REMARKS

Discussion of claim amendments.

Applicants have amended the claims as follows.

Independent claim 30 has been amended to incorporate the subject matter of claim 34, directed to where the mutant alga is "of Chlamydomonas spp.", and claim 34 has been canceled.

Also, claims 31 to 33, and 35 have been canceled.

Claim 36 has been amended to depend from claim 30, instead of now canceled claim 35.

Claim 37 has been canceled.

Independent claim 38 has been amended to recite that the mutant alga is "of Chlamydomonas spp.", which is one item of the Markush group recited in claim 42, and claim 42 has been canceled.

Also, claims 39 to 41, and 43 have been canceled.

Claim 44 has been amended to depend from claim 38, instead of now canceled claim 43.

No new matter has been added by any of the amendments to the claims, and thus, the Examiner is respectfully requested to enter the amendments to the claims.

Overview.

Applicants respectfully below present arguments in support of independent claims 30 and 38 directed to mutant alga of <u>*Chlamydomonas spp.*</u> (i.e., to species of alga in the genus *Chlamydomonas*) without restricting these claims 30 and 38 to the species *Chlamydomonas reinhardtii,* and more particularly without restricting to the *Stm6* strain, as independent claim 28 is already restricted.

Claim Objection to claim 37.

The Examiner objected to claim 37 as a substantial duplicate of claim 28. In view of this objection, applicants have canceled claim 37.

Hence, the Examiner is respectfully requested to withdraw the objection to claim 37.

Claim Rejections.

Rejection of Claims 30 and 38 under 35 USC Section 112, second paragraph.

The Examiner rejected claim 30, and asserted that the term "illuminated conditions" is confusing and vague on the grounds that it is unclear what illuminated conditions means.

Applicants respectfully draw the attention of the Examiner to the fact that the term is defined in the specification at page 27, lines 12 to 15. As the Examiner will appreciate from this definition, claim 30 defines a mutant alga of *Chlamydomonas* species which is capable of hydrogen production when there is sufficient light available to it for photosynthesis to take place. Representative indications of the light intensity are given at page 27, lines 15 to 20 of the specification, but as the person ordinarily skilled in the art would appreciate, photosynthesis can take place at very low to very high light intensities. Hence, for all intents and purposes, the effect of the phrase "illuminated conditions" is to indicate that the organism is capable of hydrogen production in the light, as opposed to when it is kept deliberately in a darkened environment.

The Examiner rejected claims 30 and 38, and asserted that the recitation "HydA" is confusing as it is unclear whether it refers to a specific hydrogenase or not.

Applicants respectfully confirm that HydA is the name of a specific hydrogenase as would be understood by the person ordinarily skilled in the art.

Applicants respectfully draw the attention of the Examiner to the fact that the process of hydrogen production in the green alga *Chlamydomonas reinhardtii* is described in the specification from line 12 of page 3 to line 5 of page 4, from which it is apparent that under illuminated, anaerobic conditions the hydrogenase, HydA, located in the chloroplast stroma, catalyses the conversion of electrons and protons to hydrogen gas which is released from the cell while ATP is generated in the chloroplast. This process also is illustrated in Figure 1 and Figure 2 and the role of HydA is clearly shown. A redox-controlled regulation mechanism operates under transient light conditions to switch from linear to cyclic photosynthetic electron transport under appropriate conditions. However, the enzyme HydA is extremely sensitive to inhibition by oxygen, as noted by Melis et al. (2000), cited and described in the specification at page 4, lines 7-15, and therefore, efforts have been devoted to temporal separation of oxygen generation from the oxygen-sensitive hydrogen production process catalyzed by the chloroplast hydrogenase, HydA.

Further, as noted by Florin et al. (2001), cited and described in the specification at page 3, lines 12-24, many organisms have an enzyme capable of catalyzing the reversible reduction of protons to molecular hydrogen. There are several phylogenetically distinct groups of hydrogenase enzymes: nickel iron hydrogenases, iron hydrogenases and metal-free types. The iron hydrogenases have been found in hydrogen-producing anaerobic bacteria and protozoa, and more recently in green algae such as *C. reinhardtii* and *S. obliquus*. **In hydrogenase nomenclature, the term "Hyd" is proposed to be reserved for these enzymes, with the**

terminal letter ("HydA") distinguishing between enzymes where necessary. A second category of iron hydrogenase is composed of mostly oligomeric enzymes that interact with NAD(P) and contain domain homologous to the NuoE and NuoF subunits of complex I, and the suggested nomenclature for the catalytic subunits is HndA. The proposal for nickel ion hydrogenases involves the use of HynSL, HupSL, and so on. So the gene HydA encodes a protein which catalyses the reversible reduction of protons to molecular hydrogen and is either a monomeric iron hydrogenase or the large subunit of a dimeric iron hydrogenase (the small subunit is referred to as HydB).

There is no reason to suppose that the HydA gene differs greatly in structure across all of the species in the genus *Chlamydomonas* although, as there always is, there will be some degree of sequence difference across species within the genus. This is well understood by the person ordinarily skilled in the art, and applicants respectfully submit that, based on the claims as amended above, the Examiner cannot make a *prima facie* case that this is not so, and thus, cannot properly present a rejection of the claims as amended above.

To return to the point made by the Examiner, applicants respectfully point out that the term "a hydrogenase" clearly refers to a group of different enzymes as discussed above. Even the subset of this group, the iron hydrogenases, is a group of enzymes. While terminology has not always been used consistently, it is clear that there is standardized nomenclature now in place in which the genes encoding iron hydrogenases are the "Hyd" genes and HydA is the name of one member of the group. HydA will nevertheless have variations in sequence across species, as all genes do, as is well understood by the person ordinarily skilled in the art. Therefore, while the HydA in *C. reinhardtii* might differ slightly in sequence from the HydA gene in other hydrogen-producing *Chlamydomonas* species, it is, without doubt, the same gene in phyllogenetic and functional terms.

The Examiner rejected claims 30 and 38, and asserted that the recitation "Moc1" is confusing as it is unclear whether "Moc1" is from a specific organism or not. With regard to the objection to Moc1, applicants respectfully submit that there is substantial description of the Moc1 gene in the specification. Moc1 encodes a transcription factor homologous to human mTERF as discussed at page 37, lines 5 to 9 of the specification. Moc1 is discussed further at page 41 starting from line 6 of the specification. It is a nuclear-encoded, mitochondrial DNA-binding protein, and deletion of the activity of Moc1 results in de-regulation of the mitochondrial electron transport pathway, as discussed in the specification at page 41, line 6 to line 24. The result is inhibition of photosynthetic cyclic electron flow, and with more electrons available, there is increased hydrogen production. A comparative sequence analysis of the Moc1 gene, which was first identified in *Chlamydomonas reinhardtii* by the present inventors shows that there are striking similarities to the human mTERF protein - - the sequence alignment is given in Figure 8. Moc1 also has homologues in *Drosophila melanogaster* and sea urchin, as discussed at page 35, lines 31-34 of the specification. Additionally, nine homologues to Moc1 with mTERF domains have been identified in the genome of *Arabidopsis thaliana*.

There is a paucity of sequence data available for algae. To applicants' knowledge, the only other green alga for which sequence data is available is *Vovox Carteri* (which, like

Chlamydomonas belongs to the Order *Volvocales).* More particularly, applicant conducted an internet search using the well known BLAST (Basic Local Alignment Search Tool). The BLAST searching demonstrated that a homolog of Moc1 exists on scaffold 37 of JGI Volvox database.

Accordingly, the Examiner is respectfully requested to withdraw the rejection of claims 30 and 38 under 35 USC Section 112, second paragraph.

Rejection of Claims 30-36 and 38-44 under 35 USC Section 112, first paragraph.

The Examiner rejected claims 30 to 36 and 38 to 44 on the basis that the claims are directed to "any mutant alga from any source expressing any HydA hydrogenase having a mutation that results in reduced activity of any mitochondrial transcription factor comprising any Moc1".

Applicants respectfully note that the specific issues concerning the assertion that "any HydA" of "any Moc1" is employed have been addressed above, in the response to the rejection under 35 USC Section 112, second paragraph.

Additionally, applicants respectfully submit that the amendment above to each of independent claims 30 and 38 restricts these claims and also dependent claims 36 and 44, respectfully dependent back to independent claims 30 and 38, to algae species within the genus *Chlamydomonas.*

The genus *Chlamydomonas* is not a large and variable group, but a tightly linked phylogenetic clade in which at least the species *Chlamydomonas applanata, Chlamydomonas chlamydogama, Chlamydomonas debaryana, Chlamydomonas dorsoventralis, Chlamydomonas elliptica, Chlamydomonas eugametos, Chlamydomonas hindakii, Chlamydomonas hydra, Chlamydomonas moewusii, Chlamydomonas reinhardtii* and *Chlamydomonas texensis* are known to produce hydrogen under anaerobic conditions. Therefore, the amended claims do not claim any mutant alga from any source, but the members of a tightly linked phylogenetic clade in which the expectation would be that hydrogen production would occur by the same mechanism using the same enzyme, HydA, under the same control mechanisms.

Applicants respectfully further note that *Chlamydomonas reinhardtii* is a useful experimental model in the way that *Saccharomyces cerevisiae* and *Arabidopsis thaliana* are powerful models for dissecting basic biological processes in yeast and plants respectively. The first draft of the *Chlamydomonas* nuclear genome sequence has been made available. See attached, Dent et al., "Functional Genomics of Eukaryotic Photosyntheses Using Insertional Mutagenesis of *Chlamydomonas reinhardtii*", vol. 137, *Plant Physiology* (February, 2005), pp. 545-556. Many tools have been developed to allow for manipulation of *C. reinhardtii.* The generation of tagged insertional mutations by nuclear transformation has allowed, for example, the studies of oxygenic photosynthesis in Eukaryotes, as photosynthesis in *Chlamydomonas* is very similar to that of land plants. Plasmid, cosmid and bacterial artificial chromosome (BAC)

libraries are used to rescue nuclear mutations, and expression of specific genes can be repressed using both antisense and RNA interference technologies. See attached, Grossman et al., "*Chlamydomonas reinhardtii* at the Crossroad of Genomics", vol. 2, no. 6, *MINIREVIEW, Eukaryotic Cell* (December, 2003), pp. 1137-1150.

Additionally, there is confirmation in Grossman et al. that gene disruption is routine once sequence information is available, and further that completion of sequence information permits targeted generation of mutations (see, pages 1142 and 1143 of Grossman et al.). Thus, there is confirmation in a publication made after the priority date of the routine nature of the relevant techniques *once* (a) a discovery concerning the utility of a gene is made and (b) it is identified and characterized so as to have sequence information available. The publication nevertheless contains no teaching or suggestion of any specific finding surrounding Moc1 or hydrogen production.

The present inventors produced the *Stm6* strain. The process involved a random insertion of the plasmid pArg7.8, carrying the Arg7 gene, into the genome of the Arginine auxotrophic strain, CC 168 followed by identification of potential state transition mutants. *Stm6* was identified and found to be blocked in state 1 due to insertion of the pArg7.8 plasmid in the Moc1 gene. There was an additional insertion in a nuclear transposon (Toc1) as discussed at page 34, lines 20-28 of the specification. PCR analysis of *Stm6* and the wild type resulted in the amplification of a 1005 bp PCR product in *Stm6*. This confirmed that the insertion caused the deletion of only part of Moc1, and the remaining 512 base pairs of Moc1 remain. Figure 8 gives the protein sequence of Moc1 and an alignment with the human transcription termination factor mTERF. With this sequence information, the person ordinarily skilled in the art may employ the tools that exist for *Chlamydomonas reinhardtii* to knock out expression of the gene, without undue experimentation. For example, antisense and RNAi techniques may be employed to silence the gene. Alternatively, such site-specific mutagenesis could be used to introduce activity destroying mutations and/or an antibody to Moc1 generated, as would be well understood by the person ordinarily skilled in the art.

Further, applicants respectfully submit that it is the reduction of Moc1 activity in the *Stm6* mutant from which it derives its ability to produce greater quantities of hydrogen. The hydrogenase HydA is naturally present in the organism and therefore there is no reason that the specification should describe how to make algae expressing HydA as the Examiner appears to require. The hydrogenase is not manipulated or altered in the present invention. Rather, the Moc1 knockout or knock down induces changes in the organism which increase linear electron transport to HydA and reduce cyclic electron transport as discussed, for example, at page 41, lines 6 to 24 of the specification. The reduction or elimination of Moc1 activity by any means will achieve this end, but it is not true, as the Examiner asserts, that this is not reasonably predictable on the face of the specification because the specification does not establish the structure of Moc1. In fact, sequence information for the Moc1 gene is provided in Figure 8 and in the sequence listing. The person skilled in the art, using the tool kit available for *Chlamydomonas reinhardtii* and techniques known to the person skilled in the art, with this information, could modify the organism, without undue experimentation.

Accordingly, the Examiner is respectfully requested to withdraw the rejection of claims 30-36 and 38-44 under 35 USC Section 112, first paragraph.

CONCLUSIONS

In view of the above amendments and remarks, applicants respectfully request the Examiner to withdraw the objection to claim 37, the rejection of claims 30 and 38 under 35 USC Section 112, second paragraph, and the rejection of claims 30-36 and 38-44 under 35 USC Section 112, first paragraph.

Allowance is earnestly solicited. If the Examiner should have any questions, he is respectfully requested to telephone the undersigned to resolve any such issues, and obviate the issuance of another Office Action.

DEPOSIT ACCOUNT

Although it is believed that no fee is due, the Commissioner is authorized to charge any deficiencies of payment associated with this Communication, or to credit any overpayment, to **Deposit Account No. 13-4365**.

Respectfully submitted,

MOORE & VAN ALLEN PLLC

Date: November 21, 2007

By: Jennifer L. Skord

Jennifer L. Skord
Registration No. 30,687
Moore & Van Allen PLLC
430 Davis Drive, Suite 500
Morrisville, NC 27560-6832

Telephone: (919) 286-8000
Facsimile: (919) 286-8199

Encls.:

Dent et al., "Functional Genomics of Eukaryotic Photosyntheses Using Insertional Mutagenesis of *Chlamydomonas reinhardtii*", vol. 137, *Plant Physiology* (February, 2005), pp. 545-556

Grossman et al., "*Chlamydomonas reinhardtii* at the Crossroad of Genomics", vol. 2, no. 6, *MINIREVIEW, Eukaryotic Cell* (December, 2003), pp. 1137-1150

Functional Genomics of Eukaryotic Photosynthesis Using Insertional Mutagenesis of *Chlamydomonas reinhardtii*[1]

Rachel M. Dent, Cat M. Haglund, Brian L. Chin[2], Marilyn C. Kobayashi, and Krishna K. Niyogi*
Department of Plant and Microbial Biology, University of California, Berkeley, California 94720–3102

The unicellular green alga *Chlamydomonas reinhardtii* is a widely used model organism for studies of oxygenic photosynthesis in eukaryotes. Here we describe the development of a resource for functional genomics of photosynthesis using insertional mutagenesis of the Chlamydomonas nuclear genome. Chlamydomonas cells were transformed with either of two plasmids conferring zeocin resistance, and insertional mutants were selected in the dark on acetate-containing medium to recover light-sensitive and nonphotosynthetic mutants. The population of insertional mutants was subjected to a battery of primary and secondary phenotypic screens to identify photosynthesis-related mutants that were pigment deficient, light sensitive, nonphotosynthetic, or hypersensitive to reactive oxygen species. Approximately 9% of the insertional mutants exhibited 1 or more of these phenotypes. Molecular analysis showed that each mutant line contains an average of 1.4 insertions, and genetic analysis indicated that approximately 50% of the mutations are tagged by the transforming DNA. Flanking DNA was isolated from the mutants, and sequence data for the insertion sites in 50 mutants are presented and discussed.

As with other model organisms, the availability of genome sequence data is revolutionizing and revitalizing research into the biology of the unicellular green alga *Chlamydomonas reinhardtii* (Grossman et al., 2003; Ledford et al., 2005). Over the past four decades, many fundamental insights into the structure, function, assembly, and regulation of the photosynthetic apparatus have come from studies of Chlamydomonas, which offer several advantages for the genetic dissection of eukaryotic photosynthesis (for review, see Davies and Grossman, 1998; Hippler et al., 1998; Grossman, 2000; Dent et al., 2001; Rochaix, 2001). First and foremost, photosynthesis is fully dispensable in Chlamydomonas, as cells can grow heterotrophically in the dark using acetate as a sole carbon source. Cells grown in the dark, however, still synthesize and assemble a fully functional photosynthetic apparatus. This allows the isolation and analysis of mutants that are unable to perform photosynthesis, and light-sensitive mutants can be maintained in complete darkness. Because Chlamydomonas is predominantly maintained in a haploid form, it is not necessary to generate homozygous nuclear mutants, and mutants affecting photosynthesis can be screened immediately following mutagenesis. Chlamydomonas has an easily controlled and rapid sexual cycle (approximately 2 weeks) with the possibility of tetrad analysis, which facilitates genetic analysis. Its rapid cell-doubling time (approximately 10 h) and microbial lifestyle mean that it is easy to grow homogeneous cultures on any scale, simplifying physiological and biochemical characterization in comparison to multicellular land plants (Ledford et al., 2005). By way of example, the application of inhibitors and generators of various types of reactive oxygen species results in uniform uptake of the chemical by each cell. In land plants, multicellularity leads to differential uptake of exogenous substances based upon the distance from or method of application, and different tissue and cell types may react differently to any given chemical, making analysis of results difficult.

In spite of these differences, however, the photosynthetic apparatus of Chlamydomonas is very similar to that of land plants, making it a useful comparative system for understanding plant metabolism and photosynthesis (Gutman and Niyogi, 2004). As a member of the division Chlorophyta, Chlamydomonas is also a useful model for investigating evolutionary relationships among the green algae and thus the origins of photosynthesis in land plants.

The first draft of the Chlamydomonas nuclear genome sequence was released in January, 2003 (Grossman et al., 2003), and a complete, fully annotated version is expected in the near future. The recent accumulation of expressed sequence tag (EST) sequence data (Asamizu et al., 1999; Shrager et al., 2003) has both facilitated annotation and given some indication of the degree of accuracy that can be achieved when using bioinformatic tools to predict gene structure from assembled sequence data in this organism. The completion of the genome sequences of *Volvox carteri* and *Ostreococcus tauri* will also aid in this endeavor to identify the complete gene set of Chlamydomonas.

[1] This work was supported by the National Science Foundation (grant no. MCB-0235878).

[2] Present address: Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139.

* Corresponding author; e-mail niyogi@nature.berkeley.edu; fax 510–642–4995.

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.104.055244.

Now that the sequencing phase of Chlamydomonas genomics is nearing completion, the next step is the functional characterization of the genes. Sequence comparison and phylogenetic approaches can be used to identify putative functional homologs of genes whose functions are known in other organisms, but mutagenesis is one of the most powerful methods for assigning function to a given gene or gene family. In Chlamydomonas, insertional mutagenesis has proved to be a very useful tool in forward genetics studies, which aim to identify genes involved in a given process. Integration of exogenous DNA into the nuclear genome of Chlamydomonas occurs predominantly by nonhomologous recombination, thus leading to random gene disruption (Tam and Lefebvre, 1993). In most cases, insertional mutagenesis creates null mutations. In comparison to point mutations, insertional mutagenesis allows the isolation of sequence flanking the mutation by methods such as plasmid rescue and PCR-based techniques. Although the recent development of a detailed molecular map (Kathir et al., 2003) has made the mapping of point mutations relatively rapid in Chlamydomonas, this is still not a viable alternative for high-throughput analysis of large numbers of mutants.

Although insertional mutagenesis has been used extensively in the investigation of many areas of Chlamydomonas biology, only one study has described the use of the technique at a genomics level. Pazour and Witman (2000) reported the use of a genomic approach, involving both forward and reverse genetics, to isolate mutations affecting the outer dynein arm of Chlamydomonas flagella. This structure consists of a total of 15 proteins, thus giving some indication of the number of target genes that were involved. Mutations in the outer dynein arm result in a characteristic slow, jerky, swimming phenotype. After screening 15,000 insertional mutants for this phenotype, mutations in 7 of the 15 target genes were identified.

The paucity of studies at the genome level illustrates the need for more extensive functional genomic analyses and resources for Chlamydomonas to complement the already considerable sequence information that is available. The generation of large mutant collections has been vital in the development and use of other model plant systems such as Arabidopsis (*Arabidopsis thaliana*; Krysan et al., 1999; Tissier et al., 1999; Parinov and Sundaresan, 2000; McElver et al., 2001; Sessions et al., 2002; Alonso et al., 2003), rice (*Oryza sativa*; Jeon et al., 2000; Chen et al., 2003; Kolesnik et al., 2004; Sallaud et al., 2004), and maize (*Zea mays*; Raizada et al., 2001; May et al., 2003). Therefore, we have initiated a large-scale forward genetics project using insertional mutagenesis that aims to saturate the Chlamydomonas nuclear genome for mutations affecting photosynthesis as part of the Chlamydomonas Genome Project (Grossman et al., 2003). In this article, we describe the mutant generation and screening methods being employed in this project. As a resource to workers in the field who will be using these mutants, the phenotypic, molecular, and genetic characteristics of a subset of mutants are reported here, in addition to flanking sequence data. The whole population of phenotypically characterized mutants and a searchable sequence database will be available to the scientific community as they are generated over the next several years.

RESULTS

Generation of Insertional Mutants

To isolate insertional mutants affecting all aspects of photosynthesis in Chlamydomonas, selection of transformed cells in the dark was necessary. Although mutants incapable of photoautotrophic growth can be isolated and maintained as acetate-requiring mutants in the light, this approach does not allow the recovery of all photosynthetic mutants (Spreitzer and Mets, 1981). Very few mutants with defects in the CO_2 fixation reactions of photosynthesis, for example, can be recovered this way, because the mutants are light sensitive.

After comparison of the growth of several wild-type Chlamydomonas strains in the dark, the strain 4A+ in the 137c genetic background was selected as the parental strain for the population of insertional mutants based on its ability to grow well and remain green in the dark. Cells were transformed with either of 2 linearized plasmids, pSP124S or pMS188 (Fig. 1), containing the *ble* gene, which confers resistance to the antibiotic zeocin (bleomycin), and transformants were selected on acetate-containing medium in the dark. Transformation efficiencies using the 4A+ strain were 86.5 transformants/μg DNA for pSP124S and 115.5 transformants/μg DNA for pMS188. Both of these efficiencies are lower than those reported for these plasmids in other studies (Lumbreras et al., 1998; Schroda et al., 2002), suggesting that 4A+ may transform at lower efficiencies than cell wall-deficient strains and other strains that were used previously. Here we report data for a total of 2,000 insertional mutants generated using pSP124S and 760 using pMS188.

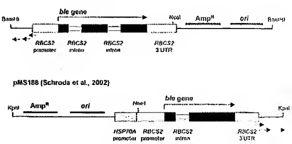


Figure 1. Diagram of linearized plasmids used for insertional mutagenesis. Relevant restriction enzyme sites are shown. Arrows indicate the approximate positions of specific primers used for TAIL-PCR.

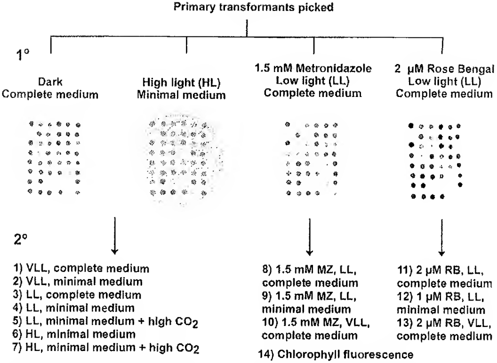


Figure 2. Outline of primary and secondary screening procedures for the isolation of mutants with photosynthesis-related phenotypes. VLL, 3 μmol photons m^{-2} s^{-1}; LL, 80 μmol photons m^{-2} s^{-1}; HL, 500 μmol photons m^{-2} s^{-1}.

Results of Phenotypic Screening

Insertional mutants were subjected to primary and secondary rounds of phenotypic screening (Fig. 2). The primary screens included incubation of the mutants at high light (HL; 500 μmol photons m^{-2} s^{-1}) on minimal medium to isolate all light-sensitive or nonphotosynthetic clones. Two generators of reactive oxygen species were used to isolate mutants that are sensitive to photooxidative stress, which often accompanies photosynthesis. Like chlorophyll, Rose Bengal (RB) generates singlet oxygen in the presence of light. By growing cells on medium containing RB, elevated levels of singlet oxygen would be present within cells and in the surrounding medium. Metronidazole (MZ), however, acts by accepting electrons from reduced ferredoxin and catalyzing superoxide formation in the chloroplast compartment of Chlamydomonas (Schmidt et al., 1977). The secondary screening methods were designed to characterize the phenotype of primary mutants more fully by assessing the degree of light sensitivity (at various light intensities) and ascertaining whether the response to generators of reactive oxygen species was dependent on photoautotrophic or heterotrophic growth conditions (Fig. 2).

The proportions of mutants in each major phenotypic class are presented in Table I. The total proportion of mutants showing a phenotype in any of the screens was 8.8%. It should be noted that the classes of mutants presented in Table I are not mutually exclusive, and thus mutants may show a phenotype in more than one of the test screens.

The largest class of mutants recovered was the acetate-requiring mutants. In agreement with Spreitzer and Mets (1981), most of these also exhibited some sensitivity to light, either at the low-light (LL; 80 μmol photons m^{-2} s^{-1}) or HL level. Secondary screening showed that 18% of the acetate-requiring mutants could be rescued, at least partially, under conditions of high CO_2. The LL-, HL-, RB-, and MZ-sensitive classes all occurred at a frequency of approximately 2.3%. Of the total number of mutants found to be sensitive to either generator of reactive oxygen species, only one-third showed sensitivity to both RB and MZ. The smallest mutant class comprised the pigment-deficient mutants, and these occurred at a frequency of 0.6%. This class included mutants that were pale green in all

Table I. *Percentage of mutants in each phenotypic class*

Phenotype	Percentage
LL sensitive ($\geq$80 μmol photons m^{-2} s^{-1})	2.3
Acetate requiring	3.8
HL sensitive ($\geq$500 μmol photons m^{-2} s^{-1})	2.3
RB sensitive	2.3
MZ sensitive	2.3
Pigment deficient	0.6

treatments, or white, yellow, or brown in at least one treatment.

Molecular Analysis of Transformants

To characterize the average number of *ble* insertion loci in each mutant, DNA gel-blot analysis was carried out on those mutants that exhibited a phenotype in any of the screens. For the population generated using the pSP124S plasmid, 85 mutants were analyzed, and 30 were analyzed for which pMS188 was the transforming plasmid. Figure 3 shows examples of the DNA gel-blot analysis. It was found that, for both plasmids, approximately 70% of the transformants contained a single *ble* insertion locus (61/85 for pSP124S and 22/29 for pMS188). The average number of *ble* insertion loci for pSP124S was 1.4 and for pMS188 it was 1.3. It should be noted that this analysis would not be able to identify clones in which multiple *ble* insertions had occurred at one locus.

In addition to probing for the sequence encoding the *ble* gene, 53 of the mutants were also analyzed for the presence of the origin of replication from the pBluescript portion of the transforming plasmid. Thirty-one of the 53 mutants (58.5%) were found to have 1 or more bands hybridizing to this sequence. Of these 31, however, 23 (74%) contained bands of the same size that hybridized to both the *ble* probe and the origin of replication probe. Because the genomic DNA was digested with *Nco*I, which should cut between these sequences in the linearized transforming plasmid (Fig. 1), bands of different sizes should be detected with the two probes. This suggests that the clones in which the same-size fragment was detected all contained tandem head-to-tail insertions at the same locus.

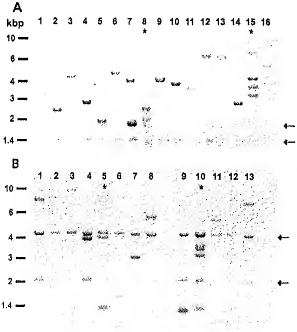


Figure 3. DNA gel-blot analysis of insertional mutants. Arrows indicate bands corresponding to endogenous *RBCS2* sequences, and asterisks indicate mutants containing multiple *ble* insertions. Size standards are shown to the left. A, Mutants generated using pSP124S. Genomic DNA was digested with *Nco*I, and the probe was a *Xba*I/*Bam*HI fragment from pSP124S. B, Mutants generated using pMS188. Genomic DNA was digested with *Nhe*I, and the probe was a *Nhe*I/*Kpn*I fragment from pMS188.

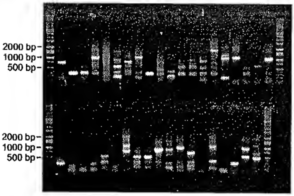


Figure 4. Agarose gel analysis of tertiary TAIL-PCR products from 39 insertional lines. Size standards are shown to the left.

Isolation and Sequencing of Flanking DNA

After secondary screening, DNA was extracted from all mutants that rescreened with the same phenotype as recorded in the primary screen. Flanking DNA was amplified from each insertional mutant line using thermal asymmetric interlaced (TAIL)-PCR (Liu et al., 1995). At least 1 DNA band was amplified in 77% of mutants where pSP124S was used as the transforming plasmid. Figure 4 shows a representative agarose gel analysis of fragments amplified from a subset of insertional mutants. The size of bands amplified using this technique ranged from <100 to 2,000 bp, with most bands being in the 100- to 1,000-bp range. Single bands were amplified in 44% of the mutants tested.

One of the problems encountered with the TAIL-PCR technique is that some of the insertion lines contained concatameric insertion events at a single locus. As insertion events included tandem arrays of the transforming DNA, sequencing of the product from TAIL-PCR only yielded plasmid sequence. For pSP124S, many of these mutants could easily be recognized by a diagnostic band of 750 bp, and several other DNA fragments also yielded only plasmid sequence. Overall, in 15.3% of the mutants from which a TAIL-PCR product was amplified, it was not possible to obtain the flanking DNA sequence due to concatamerization at the site of insertion.

Table II presents the flanking sequence results for the fragments generated by TAIL-PCR from 50 mutants. Sequences were compared to the Chlamydomonas

Table II. *Phenotypic description, flanking sequence data (by similarity to the Chlamydomonas genome sequence), and results of molecular analysis of insertional mutants*

Mutant ID	Phenotype(s)	Genome Position and Candidate Gene(s)[a]	No. of *ble* Insertions[b]
CAL005.01.01	Bleaches on HS	185: 23950–23500	n.d.
CAL005.01.13	Acetate requiring, LL sensitive	248: 49881–50798 Genie 248.7 *RBCS1* (Chlamydomonas) Genie 248.8 *RBCS2* (Chlamydomonas)	1
CAL005.01.20	RB sensitive	1725: 13500–13889	1
CAL005.01.21	LL sensitive, acetate requiring	2187: 552–627	1
CAL005.01.26	LL sensitive, acetate requiring	248: 14045–13965	1
CAL007.01.01	Bleaches at LL, RB, and MZ sensitive, low chlorophyll fluorescence	86: 72908–73019 Genie 86.13: β-7 subunit of 20S proteasome (rice) Genie 86.14: Histone-binding protein N1/N2 (*Xenopus laevis*)	1
CAL007.01.04	RB sensitive	1243: 2407–2199 Genie 1243.1 and 1243.2: Trans-splicing factor *Raa3* (Chlamydomonas)	1
CAL007.01.09	Yellow in the dark and at VLL	137: 48662–48363 Genewise 137.14.1: *crtH*; carotene isomerase (*Synechocystis* sp. PCC6803)	2
CAL007.01.11	Bleaches on HS, MZ sensitive	125: 47270–46956 Genewise 125.48.1: *hemD*, uroporphyrin III-synthase (*Synechocystis* sp. PCC6803)	
CAL007.01.17	Bleaches on MZ/HS	45: 114609–114355	1
CAL007.01.18	Slight RB sensitivity	876: 16393–16434 Genie 876.2: Acetyl CoA synthetase (Arabidopsis)	1
CAL007.01.24	LL sensitive, acetate requiring, high chlorophyll fluorescence	No genome similarity	1
CAL007.01.25	Slight RB sensitivity	1543: 7155–6541	1
CAL007.01.26	Acetate requiring, MZ sensitive	416: 20998–20655	2
CAL007.01.29	Slight acetate requirement	No genome similarity Identity to EST: 1031030D08.y1	1
CAL007.01.30	HL sensitive, acetate requiring, MZ and RB sensitive	Multiple hits, repeat region	2
CAL007.01.40	RB sensitive	387: 43424–43184	1
CAL007.01.42	RB sensitive, slight MZ sensitivity	785: 20041–19556 Genie 785.4: *HSP101* (Arabidopsis)	1
CAL007.01.43	MZ sensitive, some RB sensitivity	239: 60731–60473 Genie 239.1.1 Histone H2A (Chlamydomonas) Genie 239.37.1 Histone H3 (*Volvox carteri*) Genie 239.7.1 Histone H4 (Chlamydomonas)	1
CAL007.01.46	Slight RB sensitivity	119: 81484–81721	3
CAL007.02.02	LL sensitive, acetate requiring, partially rescued by high-CO_2, low-chlorophyll content	595: 18007–17767	1
CAL007.02.03	LL sensitive, acetate requiring	1152: 5548–5360 Genie 1152.2 ATP-synthase δ-chain (Chlamydomonas)	3-5
CAL007.02.09	RB sensitive, reduced pigment at high CO_2	91: 50935–50800 Genie 91.10: ODA1 outer dynein arm docking protein (Chlamydomonas) Genie 91.11 Silencing-related Ser-Thr kinase (rice)	1
CAL007.02.10	MZ sensitive	563: 15107–14707 Genie 563.3 Digalactosyldiacylglycerol synthase (Arabidopsis)	1
CAL007.02.18	Slight RB and MZ sensitivity	1535: 12114–12248	n.d.
CAL007.02.19	Slight RB and MZ sensitivity	851: 9886–9587 Genewise 851.5.1: Putative Cu(II)-type ascorbate-dependent monooxygenase (Arabidopsis)	n.d.
CAL007.02.21	Acetate requiring, HL sensitive, rescued by high CO_2, RB sensitive, high chlorophyll fluorescence	Multiple hits in genome, repeat region	1

(*Table continues on following page.*)

Table II. (*Continued from previous page.*)

Mutant ID	Phenotype(s)	Genome Position and Candidate Gene(s)[a]	No. of *ble* Insertions[b]
CAL007.02.27	MZ sensitive, low growth rate	**89**: 98986–99143 Genie 89.16: Potential Cu-transporting ATPase type 3 (Arabidopsis) Genie 89.17: Glutathione-requiring prostaglandin D-synthase (*Gallus gallus*)	1
CAL007.02.31	Bleaches on MZ/HS	**65**: 19631–20008 Genie 65.2: Putative replication factor (Arabidopsis)	3
CAL007.02.38	Acetate requiring	**1700**: 10826–11047 Genie 1700.6: Putative NADP oxidase (*Vibrio cholerae*) Genie 1700.2–1700.5 Histone cluster (H3, H4, H2A, H2B-IV) Genie 1700.1: Phosphoglycolate phosphatase chloroplast precursor (Chlamydomonas)	1
CAL007.02.47	Slight MZ sensitivity	**68**: 17490–17867 Genie 68.3: Protein phosphatase 2C ABI1 (Arabidopsis)	2
CAL007.03.02	RB sensitive	**1380**: 4193–4169	2
CAL007.03.03	HL sensitive, RB sensitive	Multiple hits in genome, repeat region	1
CAL007.03.08	Slight MZ sensitivity	**2640**: 5102–5626 Genie 2640.0: 70-kD heat shock protein (Chlamydomonas)	2
CAL007.03.10	MZ sensitive	Multiple hits in genome, repeat region	2
CAL007.03.21	Acetate requiring	**590**: 28275–27677 Genie 590.2 and 590.3: repair endonuclease (Arabidopsis)	1
CAL007.03.22	MZ sensitive, slight RB sensitivity	**276**: 20397–20609 Genewise.276.32.1 Dynein 11-kD light chain flagellar outer arm (Chlamydomonas) Genie 276.2 cgcr-4 protein (Chlamydomonas)	1
CAL007.03.26	MZ sensitive	Fragment 1: multiple hits, repeat region Fragment 2: **228**: 4384–4683	1
CAL007.03.32	Acetate requiring, HL sensitive, rescued by high CO_2	**3868**: 836–879	2
CAL007.03.34	HL sensitive, partially acetate requiring, rescued by high CO_2	**899**: 9585–10323	3
CAL007.03.41	Acetate requiring, RB sensitive	Fragment 1: **199**: 124–58 Genie 199.1: phospho*enol*pyruvate-dependent sugar phosphotransferase system Fragment 2: **901**: 31035–30595	4
CAL007.03.43	MZ sensitive	**248**: 42663–48811 (discontinuous) Genie 248.8: *RBCS2*	1
CAL007.03.45	Acetate requiring	**3**: 156677–156894	2
CAL007.03.46	Acetate requiring, partially rescued by high CO_2	Multiple hits, repeat region	2
CAL007.03.47	Acetate requiring, HL sensitive	**45**: 151884–152295	1
CAL010.01.02	RB sensitive	**732**: 12958–12763 Genie 732.2 Autolysin (gametolysin) precursor (Chlamydomonas)	1
CAL010.01.10	RB sensitive, pale green	**1214**: 1678–1021 (discontinuous)	1
CAL010.01.11	LL sensitive	**62**: 19180–19068 Genie 62.4 Genie 62.5 Succinate dehydrogenase (ubiquinone) iron-sulfur protein precursor (Drosophila)	1
CAL010.01.21	RB sensitive, MZ sensitive	**23**: 50197–50354	1
CAL010.01.31	RB sensitive	**102**: 2403–2506 Genewise 102.30.1: Calmodulin-binding protein (Arabidopsis)	n.d.

[a]Determined by comparison with the Chlamydomonas nuclear genome sequence, version 1.0 (http://genome.jgi-psf.org/chlre1/chlre1.home.html). Alignment of the flanking sequence with the genome sequence is indicated by scaffold number (in bold) followed by sequence range in base pairs. [b]Number of *ble* insertions determined by DNA gel-blot analysis. n.d., Not determined.

genome sequence (version 1.0; http://genome.jgi-psf.org/chlre1/chlre1.home.html) and to Chlamydomonas ESTs if no genome similarity was found. Of the 50 mutants presented, only 2 did not show similarity to any region in the genome sequence, and 1 of these showed similarity to an EST sequence. Because the integration of transforming DNA in the Chlamydomonas nucleus is sometimes accompanied by a deletion at the site of insertion, candidate genes in Table II were identified based on gene models that occur within a 10-kb interval beginning at the insertion site and extending in the direction of the *ble* insert. The identification of candidate genes was limited somewhat by incomplete assembly and annotation of the genome.

Nevertheless, likely candidate genes could be identified for several mutants (Table II). In the acetate-requiring phenotypic class, putative mutants were isolated in the Rubisco small subunit (*RBCS*) locus (see below), the ATP synthase δ-subunit gene (CAL007.02.03), and a gene involved in the phospho*enol*pyruvate-dependent sugar phosphotransferase system (CAL007.03.41). In the pigment-deficient mutant class, CAL007.01.09, which is yellow in the dark, was shown to have an insertion in a gene exhibiting homology to the *Synechocystis* sp. PCC 6803 carotene isomerase or *crtH* gene. Among mutants that are sensitive to RB and/or MZ, candidate genes include 2 heat shock protein genes, *HSP101* (CAL007.01.42) and *HSP70A* (CAL007.03.08), a putative sigma-class glutathione *S*-transferase gene (CAL007.02.27), a putative digalactosyldiacylglycerol synthase gene (CAL007.02.10), and a putative Cu(II)-type ascorbate-dependent monooxygenase gene (CAL007.02.19). Mutant CAL007.01.17 has an insertion downstream from a gene showing homology to the uroporphyrin III-synthase gene (*hemD*) from Synechocystis. This mutant is sensitive to MZ and bleaches on minimal medium in HL, suggesting that siroheme and/or vitamin B12 may be involved in the response to superoxide and photooxidative stress.

Interestingly, several mutants were found to have insertions at or close to the *RBCS* locus, which contains the *RBCS1* and *RBCS2* genes. For pSP124S (111 sequences in total), this was found in 3 independent mutants (CAL005.01.13, CAL005.01.26, and CAL007.03.43). The pSP124S plasmid contains promoter, 3'-untranslated region, and intron sequences from *RBCS2* (Fig. 2). The lines CAL005.01.13 and CAL005.01.26 both have a light-sensitive, acetate-requiring phenotype consistent with a deletion of both *RBCS* genes (Khrebtukova and Spreitzer, 1996). Genetic analysis showed that CAL005.01.13, reported previously as *dim1* (Dent et al., 2001), is tagged by the transforming DNA (Table III). Isolation of the flanking sequence from both sides of the insert by plasmid rescue showed that a deletion of approximately 36 kb of genomic DNA has occurred in CAL005.01.13, and this deletion affects the entire *RBCS* locus (Dent et al., 2001). Subsequent work with this mutant has shown that the phenotype can be rescued by complementation with either the *RBCS1* or *RBCS2* genes (R.J. Spreitzer, personal communication). The flanking sequence from CAL007.03.43 showed the insertion to be immediately downstream of the *RBCS1* gene, suggesting that the *RBCS* locus is intact in this mutant. Consistent with this analysis, the mutant does not have an acetate-requiring or light-sensitive phenotype, although it is MZ sensitive (Table II).

Table III. *Genetic analysis of insertional mutants*

Mutant ID	Recombinants[a]/Total Progeny: Progeny from Complete Tetrads	Progeny from Incomplete Tetrads	Linkage
CAL005.01.13	0/16	0/25	Yes
CAL005.01.15	11/16	14/23	No
CAL005.01.16	0/144	–	Yes
CAL005.01.21	0/20	0/43	Yes
CAL005.01.26	0/48	–	Yes
CAL005.01.28	0/16	0/31	Yes
CAL007.01.01	6/12	8/13	No
CAL007.01.08	0/40	0/15	Yes
CAL007.01.09	0/16	0/41	Yes
CAL007.01.13	2/8	4/26	No
CAL007.01.20	–	8/27	No
CAL007.01.24	9/36	–	No
CAL007.01.30	0/36	0/28	Yes
CAL007.01.39	18/40	–	No
CAL007.02.02	0/48	–	Yes
CAL007.02.05	14/44	–	No
CAL007.02.38	18/44	10/25	No

[a]Recombinants include zeocin-sensitive progeny that have the screened phenotype and zeocin-resistant progeny that lack the screened phenotype. Dash indicates no progeny of that type.

Genetic Analysis

To analyze the frequency with which the mutation is linked to the transforming DNA in the population of screened mutants, several mutants were crossed to an *mt−* wild-type strain. The progeny were then analyzed for cosegregation of the zeocin-resistance phenotype with the phenotype characterized during the screening procedure. Table III shows the linkage results of 17 crosses. A total of nine mutants (52%) showed no recombinant progeny, demonstrating linkage of the screened phenotype to the transforming DNA. With the number of progeny analyzed and assuming an average of 100 kb/cM in Chlamydomonas (Kathir et al., 2003), the transforming DNA would be inserted within 50 to 100 kb of the gene or genes resulting in the screened phenotype. More progeny would need to be analyzed to state with certainty that the mutation is indeed tagged.

DISCUSSION

The last 10 years have heralded the sequencing era in biology. As more and more genome sequences

become available, one of the most significant findings being revealed is the large number of genes for which no function is known or can be predicted by sequence similarity alone. Inactivation of a gene is generally the most direct way to understand its function. An essential tool for the functional analysis of sequenced genomes is therefore the ability to create loss-of-function mutations for all of the genes (Alonso et al., 2003). Thus far, this has only been achieved for the unicellular budding yeast *Saccharomyces cerevisiae* (Giaever et al., 2002), utilizing targeted gene replacement via homologous recombination. Unfortunately, this tool is not available in many eukaryotic organisms. Gene silencing has recently been employed to study the role of approximately 86% of the predicted genes in the *Caenorhabditis elegans* genome (Kamath et al., 2003). However, RNA interference-based methods of gene inactivation have several drawbacks, including the lack of stable heritability of a phenotype and variable levels of residual gene activity. For organisms in which homologous recombination is not available, therefore, libraries of sequence-indexed insertional mutants have many advantages (Parinov and Sundaresan, 2000). Although insertional mutagenesis has been used successfully in the generation of mutant libraries in animals (Kaiser and Goodwin, 1990; Zwaal et al., 1993; Golling et al., 2002), their strength has been most convincingly demonstrated in plants (Alonso et al., 2003). Large mutant collections exist for both T-DNA and transposon insertional lines in Arabidopsis (Sundaresan et al., 1995; Tissier et al., 1999; Sessions et al., 2002; Alonso et al., 2003), maize (May et al., 2003), and rice (Kim et al., 2004; Kolesnik et al., 2004; Sallaud et al., 2004). These banks are invaluable resources for establishing gene function in higher plants (Østergaard and Yanofsky, 2004).

To develop a resource for functional genomics of photosynthesis in Chlamydomonas, we have initiated a project to generate, screen, and obtain the flanking sequence from insertional mutants that exhibit photosynthesis-related phenotypes. This article details the phenotypic, molecular, and genetic characteristics of a subset of these mutants. Phenotypic analysis of the mutants confirmed that initial selection of transformants and subsequent maintenance of the mutants in the dark allows for the recovery of a large class of light-sensitive mutants (Table I), which might otherwise have been overlooked if nonphotosynthetic mutants were isolated by screening of light-grown cultures on minimal media (Spreitzer and Mets, 1981). Maintenance in the dark, however, may lead to the accumulation of light-sensitive, spontaneous mutations over time. This was found in the case of the CAL007.01.09 mutant (with an insertion in the carotene isomerase gene), which acquired an additional light-sensitive mutation that was revealed during genetic analysis. To minimize this problem, mutants are stored in liquid nitrogen or as a mated zygospore stock as soon as possible after isolation. RB and MZ were shown to be useful for the isolation of mutants that are sensitive to generators of reactive oxygen species. The choice of these two compounds was also found to be effective in differentiating the response to specific reactive oxygen species, as only one-third of the total number of RB- or MZ-sensitive mutants were found to be sensitive to both chemicals.

Molecular analysis of the mutant population showed that only approximately 30% of the mutants contained insertions of the *ble* gene at more than 1 locus, with an average number of insertions per clone of 1.4. This is comparable to Arabidopsis T-DNA mutant collections, in which the average number of T-DNA insertions per line is reported to be approximately 1.5 (McElver et al., 2001; Sessions et al., 2002; Alonso et al., 2003). Although a higher number of insertions per mutant means that fewer mutants are required to saturate the genome, isolation of the flanking sequence becomes more difficult when PCR-based techniques are used. In addition, the presence of numerous insertions per clone often has a negative impact on the mating ability of a clone and necessitates backcrossing to isolate the relevant mutation. It is therefore advantageous to maximize the number of clones with single inserts for both molecular and genetic reasons.

Genetic analysis showed that, in approximately 50% of the insertional mutants, the phenotype cosegregated with the transforming *ble* gene (Table III). This is in agreement with other insertional mutagenesis studies in Chlamydomonas (Niyogi et al., 1997; Fleischmann et al., 1999; Moseley et al., 2000). The tagging frequency in Chlamydomonas insertional mutagenesis therefore compares well with that reported for Arabidopsis T-DNA transformation, where as few as 35% of the mutants in a population may be tagged (McElver et al., 2001). It should also be noted that mutants in

Table IV. *TAIL-PCR cycling parameters used to isolate flanking DNA from insertional mutants*

Reaction	Step	Thermal Settings	No. of Cycles
Primary	1	95°C, 2 min	1
	2	94°C, 1 min; 62°C, 1 min; 72°C, 2.5 min	5
	3	94°C, 1 min; 25°C, 3 min; ramping to 72°C over 3 min; 72°C, 2.5 min	1
	4	94°C, 30 s; 68°C, 1 min; 72°C, 2.5 min; 94°C, 30 s; 68°C, 1 min; 72°C, 2.5 min; 94°C, 30 s; 44°C, 1 min; 72°C, 2.5 min	15
	5	72°C, 5 min	1
Secondary	1	94°C, 30 s; 64°C, 1 min; 72°C, 2.5 min; 94°C, 30 s; 64°C, 1 min; 72°C, 2.5 min; 94°C, 30 s; 44°C, 1 min; 72°C, 2.5 min	12
	2	72°C, 5 min	1
Tertiary	1	94°C, 30 s; 44°C, 1 min; 72°C, 2.5 min	20
	2	72°C, 5 min	1

which the screened photosynthesis-related phenotype is not tagged may be of interest in other fields of Chlamydomonas biology. For example, mutant CAL007.03.22 was found to contain an insertion adjacent to the gene encoding the 11-kD dynein light chain of the flagellar outer arm. It is unlikely that this would lead to the observed MZ- and RB-sensitive phenotype, but the mutant may have a linked motility phenotype that would not have been detected in our screening procedure.

Molecular analysis of the mutant population also revealed that, although all mutants analyzed had at least 1 copy of the *ble* gene, only approximately 50% of mutants had a band hybridizing to the origin of replication from the pBluescript region of the transforming plasmid. PCR screening also indicated that even fewer clones contained a fully intact origin of replication and ampicillin resistance gene (data not shown), suggesting that deletions affecting the transforming DNA occur frequently upon insertion into the Chlamydomonas genome. This illustrates why plasmid rescue has been a difficult technique to use in forward genetics studies in Chlamydomonas, as sequences required for the maintenance of the plasmid in *Escherichia coli* are frequently lost. In addition to the fact that plasmid rescue is not easily modified to higher throughput approaches, the above problem also explains why TAIL-PCR is the method that we have chosen for the isolation of the flanking sequence. Although PCR-based techniques are often difficult to optimize in Chlamydomonas due to the GC-rich nature and high occurrence of repeat regions in the genome, this article reports that TAIL-PCR was successful in amplifying fragments in almost 80% of the mutants analyzed. The only drawback of TAIL-PCR is that it cannot amplify through tandem arrays of inserts, and these occurred in approximately 15% to 20% of insertional mutants. This, however, compares favorably with Arabidopsis T-DNA mutant collections, in which 25% of left-border products and 62% of right-border products have been found to contain only T-DNA sequence (Sessions et al., 2002) using TAIL-PCR. Thus, the advantages of TAIL-PCR for higher throughput strategies outweigh its drawbacks.

Since the long-term aim of this project is to saturate the Chlamydomonas genome with mutations affecting photosynthesis, several other criteria in addition to insert number need to be examined. The number of insertional mutants required to saturate the genome is also dependent on the size of deletions that may occur at the site of insertion; larger deletions have the potential to affect multiple genes. Deletions of genomic DNA occurring at the point of insertion in Chlamydomonas range in size, but can be as large as 50 kb (Tanaka et al., 1998). The population described here appears to follow the same pattern. The mutant CAL007.01.15, for example, has a deletion of 36 kb (Dent et al., 2001), whereas CAL005.01.20 has only a few base pairs deleted at the site of insertion (data not shown). Calculations of the number of mutants needed also assume that insertion is a random event (Clarke and Carbon, 1976). Whether insertional mutagenesis is truly random in Chlamydomonas has also not been examined in previous studies. T-DNA insertion in Arabidopsis has been found to show bias against both predicted coding sequences and centromeres and to occur in preferred sites of integration or hot spots (Barakat et al., 2000; Sessions et al., 2002; Alonso et al., 2003). This work reports that, of the 50 flanking sequences isolated, 3 were found to be clustered within 50 kb of the *RBCS2* locus, and 2 were found associated with histone clusters (Table II). It is therefore possible that there is some site bias during insertional mutagenesis in Chlamydomonas, and this may be related to either sequence composition of the transforming DNA or variation in recombination frequency across the genome. It might be possible to minimize the impact of site bias in the mutant collection by using a variety of plasmids and selectable marker genes for insertional mutagenesis (Randolph-Anderson et al., 1998; Kovar et al., 2002; Depège et al., 2003). The issues of average deletion size and insertion site bias will need to be resolved once more mutants have been generated and characterized, thus allowing for a more accurate estimation of the number of mutant lines that need to be generated to achieve saturation.

Over the next several years, we aim to generate and screen 80,000 insertional mutant lines in Chlamydomonas. This will lead to the isolation of approximately 7,000 mutants affected in photosynthesis and sensitivity to photooxidative stress. Flanking sequences will be available as a searchable database within the Chlamydomonas Genome Project Web site (http://www.chlamy.org) and, when the final genome sequence is released, these sequences will be marked on the genome as an optional track within the browse function. Researchers can therefore either search the database with DNA sequences of interest or scan the genomic sequence surrounding their gene of interest for flanking sequence tags from mutants. The mutants will be available to the scientific community as mated zygospore stocks from the Chlamydomonas Genetics Center. Progeny recovered from heterozygous zygospores will represent a segregating population, which will allow for immediate genetic analysis of linkage between the mutant phenotype and the selectable marker used for transformation. Strains will also be stored frozen in liquid nitrogen to minimize the loss of mutants that are unable to mate. This population of mutants will represent the first publicly available catalogued collection of insertional mutants in Chlamydomonas, and it will be an invaluable resource for photosynthesis research.

MATERIALS AND METHODS

Media and Strains

Cultures of *Chlamydomonas reinhardtii* cells were grown heterotrophically or photoheterotrophically in Tris-acetate phosphate media (TAP) or

photoautotrophically in minimal high-salt (HS) media (Harris, 1989). Strain and mutant stocks were maintained on TAP agar medium in the dark at 25°C. For procedures that required liquid cultures, cells were grown in 50 mL TAP medium with shaking at 120 rpm either in the dark or at a very low light (VLL) intensity of 3 μmol photons m^{-2} s^{-1} at 25°C, except where otherwise stated.

The Chlamydomonas strain used to generate the population of mutants was selected for its ability to grow well and remain green in the dark on TAP medium. The growth of the standard laboratory strains CC125 (*mt*+) and CC124 (*mt*−), obtained from the Chlamydomonas Genetics Center (Duke University, Durham, NC), was compared with that of strains 4A+ and 17D−, which were obtained from J.-D. Rochaix (University of Geneva). Like CC125 and CC124, 4A+ (*mt*+) and 17D− (*mt*−) are in the 137c wild-type strain background.

For genetic analysis of the mutants generated in 4A+, a near-isogenic *mt*− strain (4A−), which showed similar sensitivity to HL and reactive oxygen species generators as 4A+, was generated by 4 backcrosses of 17D− to 4A+.

The strains used for the preparation of gamete autolysin were CC620 (137c NM subclone, *mt*+) and CC621 (137c NO subclone, *mt*−). These were also obtained from the Chlamydomonas Genetics Center.

Generation of Mutants and Genetic Crosses

Insertional mutagenesis of Chlamydomonas cells followed the transformation method of Kindle et al. (1989). One of 2 plasmids was used for transformation, pSP124S (Lumbreras et al., 1998) or pMS188 (Schroda et al., 2002), linearized with *Bam*HI or *Kpn*I, respectively (Fig. 1). Transformations with pSP124S used 1 μg plasmid DNA/transformation, whereas 0.6 μg of pMS188 were used. After transformation, the cells were allowed to recover in 10 mL TAP overnight in the dark at 25°C, with shaking at 110 rpm. The cells were then collected by centrifugation (1,300*g*, 3 min), resuspended in 300 μL TAP, and plated onto TAP agar plates containing 5 μg mL^{-1} zeocin (Invitrogen, Carlsbad, CA). The plates were maintained in the dark at 25°C for 3 to 4 weeks before the zeocin-resistant transformed colonies were picked.

Genetic crosses and tetrad analysis to assess linkage of the observed phenotype with antibiotic resistance were performed according to established methods (Harris, 1989).

Screening

Stock plates of insertional mutants were maintained in the dark on TAP agar plates. Prior to screening, the mutants were subcultured to fresh TAP plates and maintained at VLL at 25°C for 3 weeks. These VLL-acclimated mutants were used to inoculate 150 μL TAP in 96-well plates by replica plating. After 5 to 7 d of growth (VLL, 25°C), 3 μL cells were spotted onto each of the following primary screen plates: (1) TAP agar; (2) HS agar; (3) 1.5 mM MZ (Sigma, St. Louis) in TAP agar; and (4) 2 μM RB (Sigma) in TAP agar (Fig. 2). The TAP plates were maintained in the dark, the MZ and RB plates were incubated at a LL intensity of 80 μmol photons m^{-2} s^{-1}, and the HS plates were incubated at a HL intensity of 500 μmol photons m^{-2} s^{-1}. The temperature for all treatments was 25°C. All plates were scored for cell growth and bleaching after 7 to 10 d of treatment. Mutants displaying reduced growth or bleaching under any condition were picked for secondary screening.

For secondary screening, cells were grown and inoculated as described for the primary screens and submitted to 14 different treatments: (1) TAP agar (dark); (2) TAP agar (VLL); (3) TAP agar (LL); (4) HS agar (VLL); (5) HS agar (LL); (6) HS agar (HL); (7) HS agar (LL, with high CO_2); (8) HS agar (HL, with high CO_2); (9) 2 μM RB in TAP agar (VLL); (10) 2 μM RB in TAP agar (LL); (11) 1 μM RB in TAP agar (LL); (12) 1.5 mM MZ in TAP agar (VLL); (13) 1.5 mM MZ in TAP agar (LL); and (14) 1.5 mM MZ in HS agar (LL). High CO_2 atmosphere was achieved by incubating the plates in BBL GasPak CO_2 pouches (Becton-Dickinson, Franklin Lakes, NJ). In addition to these 14 treatments, the maximum chlorophyll fluorescence of the dark-grown TAP stock cultures was measured using video imaging (Polle et al., 2002).

DNA Extraction Techniques

Two different DNA extraction techniques were used in the study. For DNA required for Southern analysis, the extraction method followed that of Davies et al. (1992), excluding the final CsCl purification step. For DNA used for TAIL-PCR, cells were collected by centrifugation of 6-mL cell culture in TAP medium. The pellet was washed with 200 μL Milli-Q water, and DNA was extracted using DNAzol reagent (Invitrogen) according to the manufacturer's instructions. The final DNA pellet was resuspended in 100 μL Tris-EDTA (10 mM Tris, pH 8.0, 0.1 mM EDTA).

TAIL-PCR and Sequencing of Amplified Fragments

Genomic DNA adjacent to the insertion site of the transforming DNA was amplified using TAIL PCR (Liu et al., 1995). The method employed was optimized for Chlamydomonas. Flanking DNA was only isolated from the side of the insertion adjacent to the *ble* gene in each plasmid used, as it was found that random deletions of pBluescript sequences from the other end of the transforming DNA made amplification difficult. For pSP124S, the specific primers for primary, secondary, and tertiary reactions were RMD223 (5′-TTGGCTGCGCTCCTTCTGGCATTTAAATC-3′), RMD224 (5′-GCATT-TAAATCTCGAGGTCGAC-3′), and RMD225 (5′-GATAAGCTTGATATC-GAATTCC-3′), respectively. For pMS188, the specific primers for primary, secondary, and tertiary reactions were RMD264 (5′-GTGCTGAAGCGG-TAGCTTAGCTCC-3′), RMD255 (5′-CTCCCCGTTTCGTGCTGATCAGTC-3′), and RMD256 (5′-GAGGAGTTTTGCAATTTTGTTGG-3′), respectively. Two arbitrary degenerate primers (Wu-Scharf et al., 2000) were tested for amplification, RMD227 (5′-NTCGWGWTSCNAGC-3′) and RMD228 (5′-WGNTCWGNCANGCG-3′). RMD227 was found to amplify flanking regions successfully in most samples, whereas RMD228 only resulted in fragments in 50% of samples tested. RMD227 was therefore selected as the degenerate primer for all future reactions.

Primary TAIL-PCR reactions (20 μL) contained 1 × PCR buffer (500 mM KCl, 100 mM Tris-HCl, pH 8.3, 15 mM $MgCl_2$), 200 μM of each dNTP, 5 pmol RMD223 or RMD264, depending on the plasmid used for transformation, 60 pmol RMD227, and 2.5 units Taq polymerase (Eppendorf AG, Hamburg, Germany). The cycling parameters for all reactions of TAIL-PCR are described in Table IV.

Primary reactions were diluted 25-fold and 2-μL aliquots added directly to secondary TAIL-PCR reactions (20 μL), which contained identical components and concentrations to the primary reaction with the exception that the specific primer was replaced with RMD224 or RMD255. For the low-stringency tertiary reaction, the secondary reaction was again diluted 25-fold and either 1- or 2-μL aliquots, depending on the level of amplification achieved at the secondary stage, were added to the tertiary reaction (50 μL). Again these components were identical to that of the primary reaction, using the specific primers RMD225 or RMD256. The amplified products from both the primary and secondary reactions were analyzed by agarose gel electrophoresis. Reactions were purified as follows prior to sequencing. For samples where a single band was amplified, DNA from the tertiary PCR reaction mix was isolated using the QIAquick PCR purification kit (Qiagen, Valencia, CA). If more than one band was amplified, the fragments were separated by agarose gel electrophoresis, and individual fragments were isolated from the gel using the QIAquick gel extraction kit (Qiagen) according to the manufacturer's instructions.

For direct sequencing, 10 to 60 ng DNA were amplified with 10 pmol/reaction of RMD225 (pSP124S) or RMD256 (pMS188) using the DYEnamic ET Terminator Cycle Sequencing kit (Amersham Biosciences, Piscataway, NJ) according to the manufacturer's instructions, including the optional dilution buffer at a 1:1 (v/v) dilution. Sequencing reactions were run on an ABI3100 sequencer. Sequence data are available at the Chlamydomonas Genome Project Web site (http://www.chlamy.org).

Zygospore Storage and Cryopreservation of Cells

All mutants that showed a phenotype after the secondary round of screening were stored as both zygospores and frozen cells. For zygospore storage, each mutant was mated to the strain 4A−, and the mating mix was added to clay particles (unscented cat litter) and allowed to dry as described (Harris, 1989). Cryopreservation of cultures was carried out as described previously (Crutchfield et al., 1999).

Upon request, all novel materials described in this publication will be made available in a timely manner for noncommercial research purposes

ACKNOWLEDGMENTS

We thank Jean-David Rochaix and the Chlamydomonas Genetics Center for providing strains. We also thank Heidi Ledford for a critical reading of the manuscript.

Received October 19, 2004; returned for revision November 11, 2004; accepted November 12, 2004.

LITERATURE CITED

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen HM, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, et al (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science 301: 653–657

Asamizu E, Nakmura Y, Sato S, Fukuzawa H, Tabata S (1999) A large scale structural analysis of cDNAs in a unicellular green alga, *Chlamydomonas reinhardtii*. I. Generation of 3433 non-redundant expressed sequence tags. DNA Res 6: 369–373

Barakat A, Gallois P, Raynal M, Mestre-Ortega D, Sallaud C, Guiderdoni E, Delseny M, Bernardi G (2000) The distribution of T-DNA in the genomes of transgenic *Arabidopsis* and rice. FEBS Lett 471: 161–164

Chen SY, Jin WZ, Wang MY, Zhang F, Zhou J, Jia QJ, Wu YR, Liu FY, Wu P (2003) Distribution and characterization of over 1000 T-DNA tags in rice genome. Plant J 36: 105–113

Clarke L, Carbon J (1976) A colony bank containing synthetic Col-E-1 hybrid plasmids representative of the entire *Escherichia coli* genome. Cell 9: 91–99

Crutchfield ALM, Diller KR, Brand JJ (1999) Cryopreservation of *Chlamydomonas reinhardtii* (Chlorophyta). Eur J Phycol 34: 43–52

Davies JP, Grossman AR (1998) The use of *Chlamydomonas* (Chlorophyta: Volvocales) as a model algal system for genome studies and the elucidation of photosynthetic processes. J Phycol 34: 907–917

Davies JP, Weeks DP, Grossman AR (1992) Expression of the arylsulfatase gene from the beta 2-tubulin promoter in *Chlamydomonas reinhardtii*. Nucleic Acids Res 20: 2959–2965

Dent RM, Han M, Niyogi KK (2001) Functional genomics of plant photosynthesis in the fast lane using *Chlamydomonas reinhardtii*. Trends Plant Sci 6: 364–371

Dent RM, Kobayashi MC, Niyogi KK (2001) Functional genomics of photosynthesis using *Chlamydomonas reinhardtii*. *In* Proceedings: 12th International Congress on Photosynthesis. CSIRO Publishing, Brisbane, Australia, pp 1–6

Depège N, Bellafiore S, Rochaix J-D (2003) Role of chloroplast protein kinase Stt7 in LHCII phosphorylation and state transition in *Chlamydomonas*. Science 299: 1572–1575

Fleischmann MM, Ravanel S, Delosme R, Olive J, Zito F, Wollman F-A, Rochaix J-D (1999) Isolation and characterization of photoautotrophic mutants of *Chlamydomonas reinhardtii* deficient in state transition. J Biol Chem 274: 30987–30994

Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, et al (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. Nature 418: 387–391

Golling G, Amsterdam A, Sun ZX, Antonelli M, Maldonado E, Chen WB, Burgess S, Haldi M, Artzt K, Farrington S, et al (2002) Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development. Nat Genet 31: 135–140

Grossman AR (2000) *Chlamydomonas reinhardtii* and photosynthesis: genetics to genomics. Curr Opin Plant Biol 3: 132–137

Grossman AR, Harris EE, Hauser C, Lefebvre PA, Martinez D, Rokhsar D, Shrager J, Silflow CD, Stern D, Vallon O, et al (2003) *Chlamydomonas reinhardtii* at the crossroads of genomics. Eukaryot Cell 2: 1137–1150

Gutman BL, Niyogi KK (2004) Chlamydomonas and Arabidopsis: a dynamic duo Plant Physiol 135: 607–610

Harris EH (1989) The *Chlamydomonas* Sourcebook: A Comprehensive Guide to Biology and Laboratory Use. Academic Press, San Diego

Hippler M, Redding K, Rochaix J-D (1998) *Chlamydomonas* genetics, a tool for the study of bioenergetic pathways. Biochim Biophys Acta 1367: 1–62

Jeon JS, Lee S, Jung KH, Jun SH, Jeong DH, Lee J, Kim C, Jang S, Yang K, Nam J, et al (2000) T-DNA insertional mutagenesis for functional genomics in rice. Plant J 22: 561–570

Kaiser K, Goodwin SF (1990) "Site-selected" transposon mutagenesis of Drosophila. Proc Natl Acad Sci USA 87: 1686–1690

Kamath RS, Fraser AG, Dong Y, Poulin G, Durbin R, Gotta M, Kanapin A, Le Bot N, Moreno S, Sohrmann M, et al (2003) Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. Nature 421: 231–237

Kathir P, LaVoie M, Brazelton WJ, Haas NA, Lefebvre PA, Silflow CD (2003) Molecular map of the *Chlamydomonas reinhardtii* nuclear genome. Eukaryot Cell 2: 362–379

Khrebtukova I, Spreitzer RJ (1996) Elimination of the *Chlamydomonas* gene family that encodes the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase. Proc Natl Acad Sci USA 93: 13689–13693

Kim CM, Piao HL, Park SJ, Chon NS, Je BI, Sun BY, Park SH, Park JY, Lee EJ, Kim MJ, et al (2004) Rapid, large-scale generation of *Ds* transposant lines and analysis of the *Ds* insertion sites in rice. Plant J 39: 252–263

Kindle KL, Schnell RA, Fernandez E, Lefebvre PA (1989) Stable nuclear transformation of *Chlamydomonas* using the *Chlamydomonas* gene for nitrate reductase. J Cell Biol 109: 2589–2601

Kolesnik T, Szeverenyi I, Bachmann D, Kumar CS, Jiang S, Ramamoorthy R, Cai M, Ma ZG, Sundaresan V, Ramachandran S (2004) Establishing an efficient *Ac/Ds* tagging system in rice: large-scale analysis of Ds flanking sequences. Plant J 37: 301–314

Kovar JL, Zhang J, Funke RP, Weeks DP (2002) Molecular analysis of the acetolactate synthase gene of *Chlamydomonas reinhardtii* and development of a genetically engineered gene as a dominant selectable marker for genetic transformation. Plant J 29: 109–117

Krysan PJ, Young JC, Sussman MR (1999) T-DNA as an insertional mutagen in Arabidopsis. Plant Cell 11: 2283–2290

Ledford HK, Dent RM, Niyogi KK (2005) Chlamydomonas genomics: new approaches for a classical model system. *In* D Leister, ed, Plant Functional Genomics. Food Products Press, Binghamton, NY (in press)

Liu YG, Mitsukawa N, Oosumi T, Whittier RF (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. Plant J 8: 457–463

Lumbreras V, Stevens DR, Purton S (1998) Efficient foreign gene expression in *Chlamydomonas reinhardtii* mediated by an endogenous intron. Plant J 14: 441–447

May BP, Liu H, Vollbrecht E, Senior L, Rabinowicz PD, Roh D, Pan XK, Stein L, Freeling M, Alexander D, et al (2003) Maize-targeted mutagenesis: a knockout resource for maize. Proc Natl Acad Sci USA 100: 11541–11546

McElver J, Tzafrir I, Aux G, Rogers R, Ashby C, Smith K, Thomas C, Schetter A, Zhou Q, Cushman MA, et al (2001) Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana*. Genetics 159: 1751–1763

Moseley J, Quinn J, Eriksson M, Merchant S (2000) The *Crd1* gene encodes a putative di-iron enzyme required for photosystem I accumulation in copper deficiency and hypoxia in *Chlamydomonas reinhardtii*. EMBO J 19: 2139–2151

Niyogi KK, Björkman O, Grossman AR (1997) Chlamydomonas xanthophyll cycle mutants identified by video imaging of chlorophyll fluorescence quenching. Plant Cell 9: 1369–1380

Østergaard L, Yanofsky MF (2004) Establishing gene function by mutagenesis in *Arabidopsis thaliana*. Plant J 39: 682–696

Parinov S, Sundaresan V (2000) Functional genomics in *Arabidopsis*: large-scale insertional mutagenesis complements the genome sequencing project. Curr Opin Biotechnol 11: 157–161

Pazour GJ, Witman GB (2000) Forward and reverse genetic analysis of microtubule motors in *Chlamydomonas*. Methods 22: 285–298

Polle JEW, Kanakagiri S, Jin E, Masuda T, Melis A (2002) Truncated chlorophyll antenna size of the photosystems: a practical method to improve microalgal productivity and hydrogen production in mass culture. Int J Hydrogen Energy 27: 1257–1264

Raizada MN, Nan GL, Walbot V (2001) Somatic and germinal mobility of the RescueMu transposon in transgenic maize. Plant Cell 13: 1587–1608

Randolph-Anderson BL, Sato R, Johnson AM, Harris EH, Hauser CR, Oeda K, Ishige F, Nishio S, Gillham NW, Boynton JE (1998) Isolation and characterization of a mutant protoporphyrinogen oxidase gene from *Chlamydomonas reinhardtii* conferring resistance to pophyric herbicides. Plant Mol Biol 38: 839–859

Rochaix J-D (2001) Assembly, function, and dynamics of the photosynthetic machinery in *Chlamydomonas reinhardtii*. Plant Physiol 127: 1394–1398

Sallaud C, Gay C, Larmande P, Bes M, Piffanelli P, Piegu B, Droc G, Regad F, Bourgeois E, Meynard D, et al (2004) High throughput T-DNA insertion mutagenesis in rice: a first step towards in silico reverse genetics. Plant J 39: 450–464

Schmidt GW, Matlin KS, Chua N-H (1977) A rapid procedure for selective enrichment of photosynthetic electron transport mutants. Proc Natl Acad Sci USA 74: 610–614

Schroda M, Beck CF, Vallon O (2002) Sequence elements within an *HSP70* promoter counteract transcriptional transgene silencing in *Chlamydomonas*. Plant J 31: 445–455

Sessions A, Burke E, Presting G, Aux G, McElver J, Patton D, Dietrich B, Ho P, Bacwaden J, Ko C, et al (2002) A high-throughput Arabidopsis reverse genetics system. Plant Cell 14: 2985–2994

Shrager J, Hauser C, Chang CW, Harris EH, Davies J, McDermott J, Tamse R, Zhang ZD, Grossman AR (2003) *Chlamydomonas reinhardtii* genome project. A guide to the generation and use of the cDNA information. Plant Physiol 131: 401–408

Spreitzer RJ, Mets L (1981) Photosynthesis-deficient mutants of *Chlamydomonas reinhardtii* with associated light-sensitive phenotypes. Plant Physiol 67: 565–569

Sundaresan V, Springer P, Volpe T, Haward S, Jones JDG, Dean C, Ma H, Martienssen R (1995) Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. Genes Dev 9: 1797–1810

Tam LW, Lefebvre PA (1993) Cloning of flagellar genes in *Chlamydomonas reinhardtii* by DNA insertional mutagenesis. Genetics 135: 375–384

Tanaka A, Ito H, Tanaka R, Tanaka NK, Yoshida K, Okada K (1998) Chlorophyll *a* oxygenase (CAO) is involved in chlorophyll *b* formation from chlorophyll *a*. Proc Natl Acad Sci USA 95: 12719–12723

Tissier AF, Marillonnet S, Klimyuk V, Patel K, Torres MA, Murphy G, Jones JDG (1999) Multiple independent defective *Suppressor-mutator* transposon insertions in Arabidopsis: a tool for functional genomics. Plant Cell 11: 1841–1852

Wu-Scharf D, Jeong B-r, Zhang C, Cerutti H (2000) Transgene and transposon silencing in *Chlamydomonas reinhardtii* by a DEAH-box RNA helicase. Science 290: 1159–1162

Zwaal RR, Broeks A, Vanmeurs J, Groenen JTM, Plasterk RHA (1993) Target-selected gene inactivation in *Caenorhabditis elegans* by using a frozen transposon insertion mutant bank. Proc Natl Acad Sci USA 90: 7431–7435

EUKARYOTIC CELL, Dec. 2003, p. 1137–1150
1535-9778/03/$08.00+0 DOI: 10.1128/EC.2.6.1137–1150.2003

Vol. 2, No. 6

MINIREVIEW

Chlamydomonas reinhardtii at the Crossroads of Genomics†

Arthur R. Grossman,[1*] Elizabeth E. Harris,[2] Charles Hauser,[2] Paul A. Lefebvre,[3] Diego Martinez,[4] Dan Rokhsar,[4] Jeff Shrager,[1] Carolyn D. Silflow,[3] David Stern,[5] Olivier Vallon,[1,6] and Zhaoduo Zhang[1]

The Carnegie Institution of Washington, Department of Plant Biology, Stanford, California 94305[1]; Biology Department, Duke University, Durham, North Carolina 27708[2]; Department of Plant Biology, University of Minnesota, St. Paul, Minnesota 55108[3]; Department of Energy, Joint Genome Institute, Walnut Creek, California 94598[4]; Boyce Thompson Institute, Cornell University, Ithaca, New York 14853[5]; and CNRS, UPR 1261 Institut de Biologie Physico-Chimique, 75005 Paris, France[6]

Simple, experimentally tractable systems such *Saccharomyces cerevisiae*, *Chlamydomonas reinhardtii*, and *Arabidopsis thaliana* are powerful models for dissecting basic biological processes. The unicellular green alga *C. reinhardtii* is amenable to a diversity of genetic and molecular manipulations. This haploid organism grows rapidly in axenic cultures, on both solid and liquid medium, with a sexual cycle that can be precisely controlled. Vegetative diploids are readily selected through the use of complementing auxotrophic markers and are useful for analyses of deleterious recessive alleles. These genetic features have permitted the generation and characterization of a wealth of mutants with lesions in structural, metabolic and regulatory genes. Another important feature of *C. reinhardtii* is that it has the capacity to grow with light as a sole energy source (photoautotrophic growth) or on acetate in the dark (heterotrophically), facilitating detailed examination of genes and proteins critical for photosynthetic or respiratory function. Other important topics being studied using *C. reinhardtii*, many of which have direct application to elucidation of protein function in animal cells (26), include flagellum structure and assembly, cell wall biogenesis, gametogenesis, mating, phototaxis, and adaptive responses to light and nutrient environments (32, 44). Some of these studies are directly relevant to applied problems in biology, including the production of clean, solar-generated energy in the form of H_2, and bioremediation of heavy metal wastes.

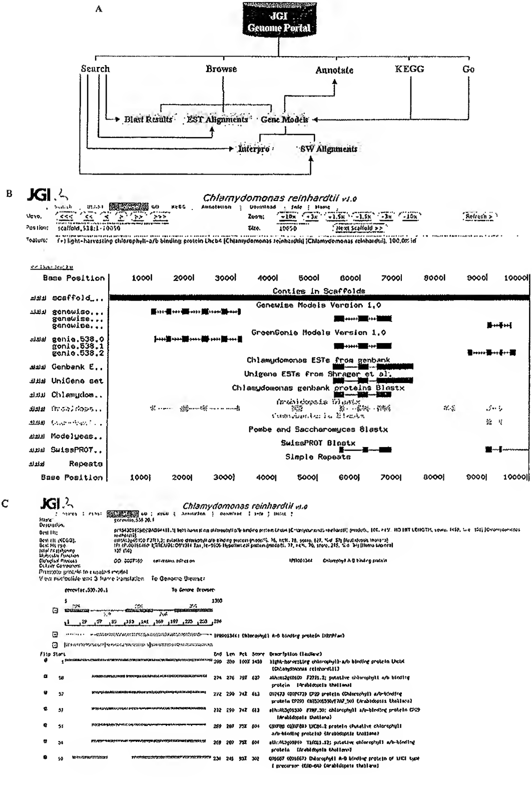
Recent years have seen the development of a molecular toolkit for *C. reinhardtii* (42, 44, 66, 98, 99). Selectable markers are available for nuclear and chloroplast transformation (4, 5, 12, 13, 30, 44, 56, 82). The *Arg7* (22) and *Nit1* (30) genes are routinely used to rescue recessive mutant phenotypes. The bacterial *ble* gene (which codes for zeocin resistance [70, 112]) is an easily scored marker for nuclear transformation, and the bacterial *aadA* gene (which codes for spectinomycin and streptomycin resistance) is a reliable marker for chloroplast transformation (39). Nuclear transformation can be achieved by particle bombardment (22, 23, 57, 73), agitation with glass beads (56, 81), or electroporation (105, 121). Generation of tagged insertional mutations by nuclear transformation has led to the rapid identification of mutant alleles (3, 17, 20, 21, 60, 108, 109, 120, 132, 138). Plasmid, cosmid (92, 139), and bacterial artificial chromosome (BAC) (66) libraries are used to rescue nuclear mutations. Expression of specific genes can be repressed using both antisense (65, 103) and RNA interference technologies (50, 58, 107; N. F. Wilson and P. A. Lefebvre, abstract presented at the 10th International Chlamydomonas Conference, 2002). In addition, endogenous transposable elements (31, 102, 127), marker rescue of *Escherichia coli* mutants (89, 136), direct rescue of *C. reinhardtii* mutants (38, 94, 132), and map-based techniques are being used to clone specific genes. Chloroplast transformation (12, 83) has permitted disruption (118) and site-specific mutagenesis of genes on the chloroplast genome (33, 34, 35, 43, 45, 46, 63, 64, 76, 129, 131, 134, 140). Reporter genes such as green fluorescent protein (36, 37), *Ars* (arylsulfatase) (19), and *Luc* (luciferase) (77; M. Fuhrmann L. Ferbitz, A. Eichler-Stahlberg, A. Hausherr, and P. Hegemann, abstract presented at the 10th International Chlamydomonas Conference, 2002) are helping to elucidate processes such as transcriptional regulation (16, 49, 87, 93, 125) and polyadenylation-mediated chloroplast RNA decay (59).

Ongoing genome projects offer the scientific community a wealth of information concerning the sequence and organization of the *C. reinhardtii* genome. Combined with the molecular toolkit, these data expand our ability to analyze gene function, organization, and evolution and to examine how environmental parameters and specific mutations alter global gene expression.

Generation of *C. reinhardtii* expressed sequence tag (EST) information was initiated in Japan (www.kazusa.or.jp/en/plant/chlamy/EST), and augmented by a National Science Foundation supported project (www.biology.duke.edu/chlamy_genome/) that has generated over 200,000 additional sequences assembled into over 10,000 "unique" cDNAs (106; unpublished data). Microarrays with representation for all of the plastid genes and approximately 3,000 nuclear genes (48, 68) have been used to probe global gene expression in wild-type (48, 68) and mutant strains (Z. Zhang and A. R. Grossman, unpublished results). Furthermore,

* Corresponding author. Mailing address: The Carnegie Institution of Washington, Department of Plant Biology, 260 Panama St., Stanford, CA 94305. Phone: (650) 325-1521. Fax: (650) 325-6857. E-mail: arthur@andrew2.stanford.edu.

† Carnegie Institution of Washington publication no. 1627.


A
JGI
Genome Portal
Search
Browse
Annotate
KEGG
Go
Blast Results
EST Alignments
Gene Models
Interpro
SW Alignments
B
JGI
Chlamydomonas reinhardtii v1.0
Base Position
Contigs in Scaffolds
Genewise Models Version 1.0
GreenGenie Models Version 1.0
Chlamydomonas ESTs from genbank
Unigene ESTs from Shrager et al.
Chlamydomonas genbank proteins Blastx
Pombe and Saccharomyces Blastx
SwissPROT Blastx
Simple Repeats
C
JGI
Chlamydomonas reinhardtii v1.0

the genomic information has aided in the generation of tools for map-based cloning, based on linkage of genetic and physical markers (55, 126).

The accumulation of cDNA sequence information and development of robust molecular markers has stimulated the interests of the Joint Genome Institute (JGI), Department of Energy, and under the leadership of one of us (D. Rohksar), a rough draft of the near-complete genome sequence was made publicly accessible in the early part of 2003. This sequence has been partially annotated and both cDNA information and molecular markers have been anchored to the sequence. These advances have dramatically enhanced the utility of *C. reinhardtii* as a model system.

NUCLEAR GENOME SEQUENCE

Assembly and annotation of the genome. The nuclear genome of *C. reinhardtii* is 100 to 110 million bp, comprising 17 genetic linkage groups (55), with a very high GC content (nearly 65%) that results in cloning difficulties and limits the length of reads from shotgun sequencing reactions. Generating a high-quality genome sequence has therefore presented unusual challenges. Sequencing strategies being used involve production of random genomic fragments of ~3 and ~6 kbp, cloning of the fragments into plasmids, and obtaining paired end sequences of the insert DNA. Paired end sequences from 35 to 40 kbp fragments in fosmid vectors are also being generated. This information is being integrated with end sequence data from 15,000 BAC clones (see "Alignment of Genetic and Physical Maps").

With a sequence redundancy of nearly 10-fold, the randomly sequenced fragments generated by the strategies described above can be assembled into "contigs" (contiguous stretches of reconstructed sequence obtained from overlapping end sequences) that are further linked together into "scaffolds" (longer stretches of reconstructed sequence interrupted by "gaps" whose size is roughly known based on spanning clones). A preliminary rough draft of the *C. reinhardtii* genome is already available at the JGI Chlamydomonas Web site (see below). A high-quality draft genome assembly is anticipated by the fall of 2004.

We plan to generate a complete sequence reconstruction of *C. reinhardtii* chromosomes by linking together sequence scaffolds using genetic and clone-based physical maps (see "Alignment of Genetic and Physical Maps"). Sequencing of selected regions of the genome is likely to be finished by further targeted efforts to close gaps in scaffolds and by resequencing low-quality regions to achieve a uniform error rate of less than one error per 10,000 bases. The ultimate goal is a high-quality reference genome sequence.

Annotation of the gene content of *C. reinhardtii* is being facilitated by copious EST information produced by several projects (see below) and availability of modern gene-finding methods that exploit expressed sequence evidence, statistical signatures of coding regions, and conservation of deduced polypeptide sequences with known proteins from other organisms. One intriguing possibility for further analysis of the *C. reinhardtii* genomic sequence is to compare it with sequence information from the colonial alga *Volvox carteri*, with the goal of highlighting coding regions that may be unique to the chlorophyte algae, and possibly to identify putative conserved regulatory regions. While computational methods can only reliably predict coding regions, the large scale EST collections will enable many 5′ and 3′ untranslated regions (UTRs) to be directly determined. Furthermore, probes synthesized based on ab initio gene predictions can be used to identify and clone rare transcripts. Integration of complementary community informatics resources centered on the genome will provide a comprehensive view of the *C. reinhardtii* genome that is readily accessed by many different network locations (see "Toward an Integrated Database").

Chlamydomonas Genome Portal. Genomic information generated at JGI can be accessed through the JGI Chlamydomonas Genome Portal (www.jgi.doe.gov/chlamy), which is intended as an archival, Web-based source for *C. reinhardtii* genomic sequence information and associated annotations (Fig. 1A). Prior to initial publication of the genome sequence and its annotation and analysis, items presented on the JGI site should be considered to be preliminary results and a community resource.

Various precalculated features identified on the genome (exons; genes; mRNA, EST, or unigene alignments; markers for mapping; protein BLAST hits; etc.) are organized in "tracks" using a graphical interface similar to that developed at Santa Cruz for the human genome (54) (Fig. 1B). Clicking on (selecting) a predicted gene will display a page (Fig. 1C) showing protein and transcript sequences, precalculated BLAST results (1), and InterPro (79) determinations of protein domains. Clicking on an EST, unigene, or mRNA alignment displays a graphical view of the alignment as well as information at the sequence level and BLAST results relative to known proteins from various organisms.

Users can reach a genomic region of interest in a variety of ways. One can perform BLAST analysis against the genome and view resulting alignments in the context of all the other database features. For example, comparing an *Arabidopsis* protein to the *C. reinhardtii* genome with BLAST would access the region of the *C. reinhardtii* genome with a similar sequence, immediately recovering the gene at that location. There are tracks for predicted gene structures based on the GeneWise (9) and GreenGenie (Susan Dutcher, personal communica-

FIG. 1. (A) Schematic of JGI genome portal. The diagram shows the internal connections of the JGI Genome Portal. Information on BLAST results, EST alignments, and gene models can be accessed through the Search page. From the gene model information page, or protein page, InterPro domains and Smith-Waterman alignments to protein databases are displayed with a graphical interface. With the version 2.0 release GO and KEGG will be available, as well as the ability to annotate gene models. The Chlamydomonas Genome Portal is accessible at www.jgi.doe.gov/chlamy. (B) Browse view. Screen shot of the browse view for several gene models displayed on the genome. Displayed simultaneously are overlapping EST alignments and Blastx results. (C) Protein page. The protein page displays information about a gene model. InterPro results, Smith-Waterman alignments, and the protein and transcript sequence for this model can be retrieved from this page.

TABLE 1. cDNA libraries

Library	Conditions	Strain	Normalization	Project no.	No. of clones
Core	TAP light, TAP dark, HS + CO_2, HS	21gr	Not normalized	874	768
Core	TAP light, TAP dark, HS + CO_2, HS	21gr	Normalized	894	10,080
Core	TAP light, TAP dark, HS + CO_2, HS	21gr	Subtracted (894)	1024	12,096
Stress I	NO_3 to NH_4 (30 min, 1 and 4 h), NH_4 to NO_3 (30 min, 1 and 4 h), TAP-N (30 min, 1 and 4 h), TAP-S (30 min, 1 and 4 h), TAP-P (4, 12, and 24 h)	21gr	Normalized	963	12,000
Stress II	NH_4 to NO_3 (24 h), H_2 production (0, 12, and 24 h), TAP + H_2O_2 (1, 12, and 24 h), TAP + sorbitol (1, 2, 6, and 24 h), TAP + Cd (1, 2, 6, and 24 h)	21gr	Normalized	1031	10,752
Deflagellation	15, 30, and 60 min	21gr	Normalized	1030	12,480
S1D2		S1D2	Normalized	925	124
Gamete	2, 8, 10, 12, 15, and 17 h	21gr	Normalized	1112	—[a]
Zygote	30 and 60 min				
Stress III	TAP-Fe, TAP-Cu, TAP-O_2, TAP high light, HS high light	21gr	Normalized	3510	—

[a] —, ongoing.

tion) algorithms, as well as for alignments of publicly available ESTs (106), molecular markers (55), array elements, and known protein sequences from specific organisms.

Since a BLAST analysis of the genome against all proteins in GenBank has already been performed and will be periodically updated, one can text search through the names of precomputed alignments. Other access points include the GO (gene ontology) and KEGG (Kyoto Encyclopedia of Genes and Genomes) links that organize genes into functional groupings.

The JGI Chlamydomonas Genome Portal is in a dynamic state of development. Assignment of gene functions is a feature of any genome project that is continually being informed by sequence similarities, experimental evidence, phylogenetic data, and expression profiles. To capture the richest annotation of the *C. reinhardtii* genome, the JGI portal includes interfaces for community annotation, allowing experts around the world to add their input, and incorporates links to publications, experiments, and descriptive text. New features being integrated into the JGI Portal include tracks showing spanning BACs and fosmids. Improved gene models will be merged ab initio with EST/mRNA evidence, increasing the number of complete gene predictions (including UTRs) and revealing alternatively spliced transcripts. Sequence signals for transmembrane spanning regions, signal peptides, and targeting sequences will also be computed and added to the site. Linkages to and from JGI pages to other community resources, notably ChlamyDB, are being developed, as described below under "Toward an Integrated Database."

THE TRANSCRIPTOME

Efforts are currently under way to identify transcribed regions of the genome and to analyze their expression patterns.

cDNA information. After a pilot experiment by S. Purton, a collection of 37,940 5'-end ESTs was generated for *C. reinhardtii* by the Kazusa DNA Research Institute in Japan (2). Normalized, size-selected libraries were generated from cells grown under low- or high-CO_2 conditions. A National Science Foundation-supported cDNA project performed at the Carnegie Institution of Washington and the Genome Technology Center at Stanford has led to the generation of cDNA libraries constructed from RNA isolated from cells exposed to a variety of different conditions (Table 1); these libraries were normalized prior to sequencing individual clones. One library is from the field isolate S1D2 (41), which has numerous sequence polymorphisms but is interfertile with the laboratory strain 21gr, and is used for map-based cloning of mutant alleles (55). Nearly 200,000 clones have been sequenced from their 3' and 5' ends (106), and full-length sequences are being generated. Our assembly protocol is based on the commonly used Phrap program, which takes into account sequence quality. The assembly generates assemblies of contiguous ESTs (ACEs), which theoretically represent unique genes (106; J. Shrager, C.-W. Chang, J. Davies, E. H. Harris, C. Hauser, R. Tamse, R. Surzycki, M. Gurjal, Z. Zhang, and A. R. Grossman, presented at the proceedings of the 12th International Congress on Photosynthesis, 2001) (www.biology.duke.edu/chlamy/PDF/Shrager2003.pdf). Sequences from the ~10,000 ACEs in the assembly designated 20021010 (dated 10 October 2002) have been annotated on the basis of BlastX homology to potential homologs in other organisms. We are currently preparing a final assembly of all of the EST data, which will include those from S1D2 as well as from the Purton and Kazusa projects. Knowing the distribution of ESTs among the cDNA libraries and the conditions used for library generation, we can infer a qualitative image of the expression pattern of specific genes. Accordingly, we have identified several genes represented by multiple cDNAs in the stress libraries (including arylsulfatase, phosphatases, and regulatory proteins) that are not represented in the core library.

Microarray construction and application. The DNA microarray is currently the most commonly used and widely applicable technique for the global analysis of gene expression. We have completed and are distributing a first generation cDNA array. A region of each cDNA 3' end was amplified using a universal primer in the vector and a specific primer ~400 bp upstream of the 3' end. PCR products were purified and printed onto GAPS II amino silane-coated slides (Corn-

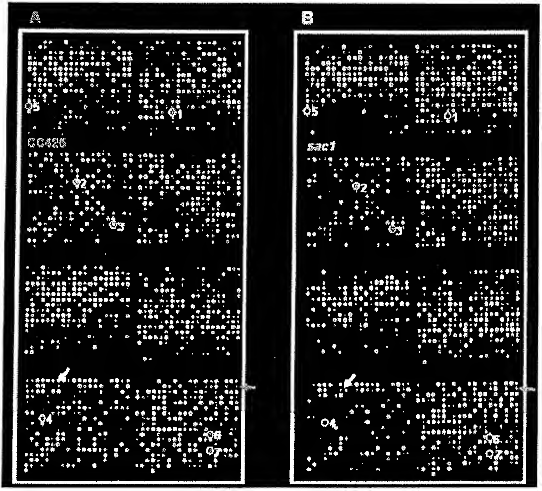


FIG. 2. *Chlamydomonas* microarrays. Shown are microarray images generated after 24 h of sulfur starvation of the parental strain (CC-425, left panel) and the *sac1* mutant in the same genetic background (CC-3794, right panel). Red fluorescence indicates an increase in the level of the transcript during sulfur deprivation, while green fluorescence indicates a decrease in transcription. Spot 1 represents the *Ecp76* gene (963017H04), and spot 6 represents an LI818 gene (894097E05), which encodes a polypeptide that is part of the light-harvesting protein family. The functions of the genes represented by spots 2 to 5 and 7 are not known (all of these spots are circled). The orange arrow marks a gene (encoding a putative poly(A) binding protein; 894006E07) for which the transcript increases in both CC-425 and the *sac1* mutant, while the white arrow marks a gene whose transcript increases in *sac1* but not CC-425 cells.

ing), with each slide carrying four replicate spots of each cDNA fragment. For version 1.0, we chose clones with high-quality sequence information from 2,761 distinct ACEs. As of January 2003, a slightly different version is being distributed (version 1.1), with ~300 additional genes amplified either from our EST libraries, or from other sources; many were kindly provided by other laboratories. Within 2 years we plan to generate an array representing the entire *C. reinhardtii* genome.

We and others have already performed experiments with these arrays. Recently, we have identified genes activated by high-intensity light under low-CO_2 conditions (48); these genes encode photorespiratory proteins, proteins that combat the accumulation of toxic oxygen radicals, polypeptides that function in concentrating inorganic carbon and several proteins of unknown function. Expression studies have also been performed with wild-type and mutant cells transferred from nutrient-replete to sulfur-deficient medium. For example, the *SacI* gene controls the acclimation of cells to sulfur deprivation conditions and encodes a regulatory protein (17, 18) that has some similarity to transporters with 12 membrane-spanning helices. Figure 2 shows a set of microarrays generated for CC-425 and the *sac1* mutant following imposition of sulfur deprivation. A number of transcripts were found to increase dramatically during starvation. Some encode proteins involved in sulfur metabolism (e.g., the *Ars* gene [which encodes aryl-

sulfatase] and the *Ats1* gene [which encodes ATP sulfurylase]) or other cellular processes (e.g., *Ecp76*, which encodes a cell wall polypeptide specific to sulfur stress cells [116]), while the functions of several others remain unknown (Fig. 2).

Similar studies are being conducted (in the Grossman laboratory), on phosphorus and nitrogen limitation, as well as on the physiological effects of different light qualities. Other microarray studies have been initiated with Krishna Niyogi (high-light-activated genes), Donald Weeks (CO_2-activated genes), and Jean-David Rochaix (mutants in photosynthetic function). We have also distributed several hundred arrays to researchers working on *C. reinhardtii*, and it is expected that a large corpus of data will be generated in the coming months that should begin to reveal global and interacting regulatory features of the genome. A specific microarray section is being introduced into the Chlamydomonas Genome Project Database in which all relevant information regarding array elements (sequence, position on the array, ACE and gene models and their annotation etc.) will be made available.

ALIGNMENT OF GENETIC AND PHYSICAL MAPS

An important component of the genome project has been the placement of molecular markers onto the *C. reinhardtii* genetic map, with the aim of facilitating map-based cloning of genes identified by mutations. Over the last 50 years, more than 200 phenotypic markers (mostly mutations) have been mapped onto the 17 *C. reinhardtii* linkage groups, and recently, more than 270 molecular markers have been placed on the linkage map. Some of these have been correlated with mutant data, allowing for the alignment of the physical and genetic maps. The defined physical markers are either restriction fragment length polymorphism- or PCR-based markers. The positioning of these markers onto linkage groups provides, on average, a map in which any given point on the *C. reinhardtii* genome is within 2 cM of a mapped molecular marker (55, 126), corresponding to 150 to 200 kbp of genomic sequence.

To facilitate the use of the molecular map for map-based cloning, a BAC library of more than 15,000 clones has been generated and arrayed, providing an eightfold coverage of the nuclear genome. (Individual BAC clones or the entire library can be obtained from the Clemson University Genomics Institute: www.genome.clemson.edu). JGI has sequenced both ends of all clones in this library, and this information is available and can be searched using BLAST on the JGI Web site (bahama.jgi-psf.org/prod/bin/chlamy/home.chlamy.cgi). More than 2,500 of these clones, focusing on those containing mapped molecular markers, have been fingerprinted and placed into overlapping BAC contigs. The BAC contigs now cover more than 25% of the genome. As the assembly of the nuclear genome proceeds, by linking together sequence scaffolds, it will be increasingly useful to compare BAC end sequences with the genomic sequence to place additional BACs onto the physical/genetic map. Ultimately, a tiling path of BAC clones corresponding to the complete *C. reinhardtii* genetic and physical maps will be generated.

The information already available has made it possible to apply map-based cloning strategies to the identification of mutant alleles in *C. reinhardtii*, e.g., *lf1* (R. Nguyen and P. Lefebvre, presented at the he 10th International Conference on the Cell and Molecular Biology of Chlamydomonas, 2002) and *bld2* (27). The *Bld2* gene was cloned by identifying overlapping BAC clones covering 720 kbp of genomic sequence corresponding to 4.5 cM on linkage group III. The BAC clone containing the wild-type *Bld2* gene was identified by transforming individual BAC clones into *bld2* mutant cells to rescue the mutant phenotype.

Map-based cloning will be greatly accelerated by a high density of genetically mapped polymorphisms between the laboratory strain 21gr and field isolate S1C5, which is very similar to S1D2. Sequence information already available suggests that the frequency of polymorphisms between the laboratory and wild-isolate strains is surprisingly high. In a survey of more than 29,000 nucleotides from the 3' UTR of 62 transcripts, there were 2.7 base substitutions and 0.54 insertions or deletions per 100 bases. This level of sequence polymorphism will allow any new mutation in a laboratory strain to be mapped both genetically and physically. A protocol for mapping any new mutation by crosses to S1C5 followed by PCR-based detection of a set of molecular markers was recently described (55). Once a mutation has been mapped to a genetic interval, more detailed fine-structure mapping may require that additional molecular markers in the interval of interest be identified. Such markers can be easily obtained from DNA sequence in regions of interest by searching for microsatellite sequences [usually (GT)n repeats]. Thousands of microsatellites, dispersed throughout the genome, can be converted into PCR-based molecular markers by designing specific oligonucleotide primers for PCR amplification of the microsatellite-containing sequence, followed by identification of the different alleles by sizing products on gels (the different alleles will have different numbers of GT repeats). Kang and Fawley (52) have used this procedure to map microsatellite sequences in *C. reinhardtii*.

ORGANELLE GENOMES

A complete *C. reinhardtii* mitochondrial genome sequence is available (GenBank accession U03843). This 15.7-kb genome encodes the cytochrome *b* and cytochrome oxidase apoproteins, six NAD dehydrogenase subunits, a protein resembling reverse transcriptase, large and small mitochondrial rRNAs (fragmented), and three tRNAs (GenBank accession U03843). All other mitochondrial components are presumably encoded in the nuclear genome.

Completion of the entire sequence of the chloroplast genome of *C. reinhardtii* has permitted the generation of mutations in all of the genes on that genome (except where the lesions are lethal) and an analysis of transcripts that emanate from different genomic regions. The complete sequence has also enabled the production of a chloroplast genome microarray that can be used for analyzing the global accumulation of chloroplast transcripts under different environmental conditions.

Chloroplast genes and their expression. The *C. reinhardtii* chloroplast genome is 203.8 kbp (GenBank accession number BK000554) and contains 99 genes, including 5 rRNA genes, 17 ribosomal protein genes, 30 tRNAs specifying all of the amino acids, and 5 genes encoding the catalytic core of a eubacterial-type RNA polymerase (72). Figure 3 depicts the circular genome, its known genes, and the positions of those that have

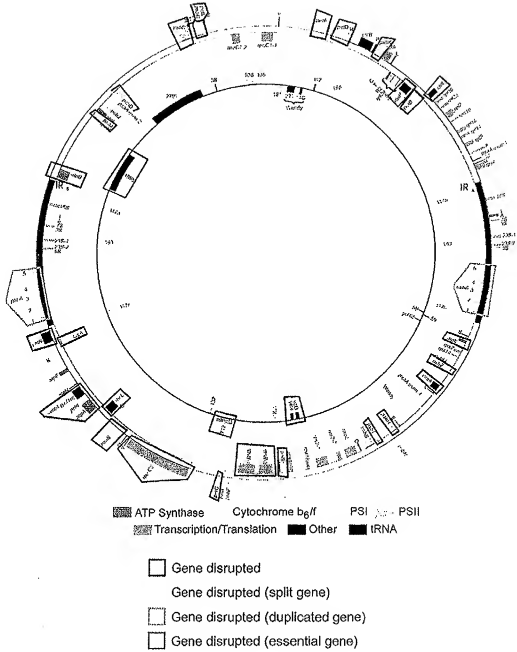


FIG. 3. Chloroplast genome. The *C. reinhardtii* chloroplast genome and its genes are shown. Those that have been disrupted are highlighted.

been disrupted. The genome contains a staggering number of small dispersed repeats (SDRs) that mostly populate intergenic regions.

The structure and gene content of the *C. reinhardtii* chloroplast chromosome are conventional, with a ribosomal DNA-containing inverted repeat separating two single copy regions. When compared to the chloroplast DNA (cpDNA) of land plants, the *C. reinhardtii* genome has a few noteworthy fea-

TABLE 2. Genes disrupted on the chloroplast genome of *C. reinhardtii*

Gene or ORF	Function[a]	Essential	Reference(s)
psaA	PSI	No	95
psaB	PSI	No	95
psaC	PSI	No	117
psaJ	PSI	No	33
tscA	PSI	No	40
ycf3	PSI	No	10
ycf4	PSI	No	10
psbA	PSII	No	7
psbC	PSII	No	100
psbD	PSII	No	29
psbE	PSII	No	78
psbH	PSII	No	85, 113
psbI	PSII	No	61
psbK	PSII	No	118
psbT	PSII	No	86
psbZ	PSII	No	115
petA	Cytochrome b_6f	No	62
petB	Cytochrome b_6f	No	62
petD	Cytochrome b_6f	No	62
petG	Cytochrome b_6f	No	8
petL	Cytochrome b_6f	No	119
atpA	ATP synthase	No	25
atpB	ATP synthase	No	104
atpE	ATP synthase	No	96
ccsA	Heme attachment	No	133
cemA	Envelope transporter	No	101
chlL	Chlorophyll synthesis	No	114
chlN	Chlorophyll synthesis	No	14
clpP	Protease	Yes	47, 71
ORF1995	Unknown	Yes	11
rbcL	Rubisco	No	110
rpoB1	Transcription	Yes	35
rpoB2	Transcription	Yes	35
rpoC2	Transcription	Yes	35
rps3	Translation	Yes	69

[a] PSI, photosystem I; PSII, photosystem II.

tures: (i) an unusual gene, *tscA*, that encodes an RNA that is involved in *trans*-splicing of *psaA* transcriptional segments; (ii) a split *rpoC1* gene; (iii) the presence of *tufA*, which encodes elongation factor Ef-Tu; (iv) two large open reading frames (ORFs) (1,995 and 2,971) of unknown but essential function; and (v) an absence of *ndh* genes, which encode polypeptides critical for chlororespiration, a process first reported in *C. reinhardtii* (6). The *ndh* genes are ubiquitous on land plant cpDNA.

Gene disruption is routine for *C. reinhardtii* chloroplast genes, and even the so-called essential genes can be functionally analyzed by weakening their translation initiation codons (71). The completion of the genome sequence does not offer many new gene candidates for functional analyses but does provide landmarks necessary for gene manipulation and the analysis of global plastid gene expression. Table 2 lists genes marked in Fig. 3 as having been disrupted; the total is an impressive 35 genes in which only 6 could not be brought to homoplasmicity.

The analysis of the chloroplast genome enables researchers to define previously undiscovered genes and to measure expression of known genes. Sequence alone does not necessarily presage identification of a full genomic complement, and some

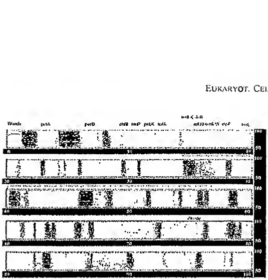


FIG. 4. SDR sequences on the chloroplast chromosome. The first 100 kb of the chloroplast chromosome were analyzed for SDRs using a genome self-comparison with the program Pipmaker (bio.cse.psu.edu/cgi-bin/pipmaker?basic). The approximate locations of genes are shown on the top row only; the "Wendy" transposon and its disabled duplicate copy are shown on the top row at around position 75,000. The thin gray line represents one copy of the large inverted repeat. Each dot represents a repeat of the sequence along the top line; e.g., Wendy is duplicated, so a second line appears underneath it. The SDRs are represented by the large numbers of dots, whose sequence identity to the particular place on the genome ranges from 50 to 100% as shown in the scale on the right.

genes (like *tscA*) may not encode proteins. To complicate matters, three of the four major photosynthetic complexes (photosystem I, photosystem II, and the cytochrome b_6f complex) contain small chloroplast-encoded polypeptides with ORF sizes that would frequently arise by chance in the genome. For this reason, annotation of the ORFs was limited to those at least 100 residues long. Since small genes or non-protein-encoding genes should nonetheless be represented in the transcript pool, a comprehensive RNA filter blot analysis was undertaken, using RNA isolated from cells grown under a range of environmental conditions. As reported by Lilly et al. (68) the accumulation of chloroplast transcripts is strongly affected by culture conditions. Under conditions in which most investigators grow their cells—in rich medium and under continuous light—chloroplast transcript accumulation is relatively high. This is consistent with the observations that substantial decreases in the cpRNA content do not, in the short term, visibly affect the synthesis of most chloroplast polypeptides (28). Under conditions of abiotic stress, changes in transcript accumulation range from subtle to as much as eightfold. Increases in the levels of some transcripts in response to phosphate deprivation appear to be mediated, at least in part, by polynucleotide phosphorylase (Y. Komine and D. Stern, unpublished results), a nuclear-encoded, chloroplast RNase whose activity is modulated by physiologically relevant phosphate concentrations (135).

SDRs. The SDRs that have colonized intergenic regions of the cpDNA (Fig. 4) present a fascinating evolutionary puzzle. Of sequenced cpDNAs within the chlorophytes, which include land plants as well as green algae, only *Chlorella* sp. appears to have numerous SDRs (72). Surprisingly, there is almost no sequence similarity between the SDRs of *Chlorella* and *C. reinhardtii*, suggesting that SDR amplification might share a common mechanism but be sequence independent. The rela-

tively balanced distribution of SDRs in the *C. reinhardtii* chloroplast genome raises questions concerning both their origin and function. Did an ancient invasion of a transposable element subsequently lead to the dispersal of smaller fragments, or did a nuclear mutation somehow permit or foment accumulation of SDRs? It has been suggested (15) that short repeats may be associated with rearrangement of chloroplast genes or that they might function as binding sites for proteins that participate in gene expression. Interestingly, SDR-rich sequences upstream of *petA* exhibit a conformational (torsional) response to light, which is correlated with increased transcriptional activity (122).

In summary, chloroplast genomics in *C. reinhardtii* has provided sophisticated tools for analyzing and manipulating cpDNA and has raised fascinating evolutionary questions. Recent years have seen accelerated cloning and analysis of nuclear genes encoding chloroplast regulatory factors (97, 99), which will stimulate studies on their interactions with chloroplast mRNAs and with one another (24, 137). With the sequencing of the *C. reinhardtii* nuclear genome, whole new families of putative regulators of chloroplast gene expression will emerge, presenting an opportunity to build an integrated image of genetic interactions between the nuclear and chloroplast genomes and how they are fine-tuned by critical features of the environment.

TOWARD AN INTEGRATED DATABASE

Use of available databases. One strength of *C. reinhardtii* as a model system lies in the extent to which it has been used for genetic and physiological characterization of biological processes. With the advent of *C. reinhardtii* genomics, we are poised to link phenotypes, alleles, and expression and sequence features into an integrated database.

The major goals of database construction are to (i) provide user-friendly points of access for the sequence data, (ii) connect genomic features to the classical biology of the organism, (iii) provide tools for viewing and querying genomic and gene expression data, and (iv) generate resources and tools for cross-species comparisons as data from related algal species become available.

Currently the genomic and organismal data are dispersed among three databases: (i) ChlamyDB, which contains information on genetic loci, mutant alleles, and sequenced genes, descriptions of strains, bibliographical citations, and community member information; (ii) ChlamyEST, which contains sequence data (EST, contigs, unigene, chloroplast, mitochondria) and gene annotations; and (iii) the JGI Chlamydomonas Genome Portal (see "Chlamydomonas Genome Portal" above), which contains the nuclear genome sequence, gene model predictions, and preliminary annotation data. All three databases are accessible through search engines, and both the Chlamydomonas Genome Project and the JGI Web sites include on-line Blast utilities, with additional specialized datasets available at ChlamyEST containing sequences from the *Volvocales* (including *Chlamydomonas*, *Volvox*, *Eudorina*, *Pandorina*, *Dunaliella*, and *Haematococcus*, among others) and BAC end sequences.

Integration of the databases. (i) Unification of ChlamyDB and ChlamyEST. The near-term challenge is to link all *C. reinhardtii*-related data sets in a seamless manner. To this end we will unify data maintained in ChlamyDB and ChlamyEST and establish links between this unified database and the JGI Chlamydomonas Genome Portal. The Chlamydomonas Genome Project is implementing a version of the Generic Model Organism Database (111) with the aim of integrating genetic, sequence, and bibliographic information. Figure 5 presents a schematic of the proposed unifications. At the core of this project is the underlying "chado" database schema, designed to integrate the *Drosophila melanogaster* data in FlyBase into distinct modular components with tightly defined dependencies ("Sequence," which contains biological sequences and annotation; "Genetics," which houses alleles and relationships between alleles and phenotypes; "Map," which contains any type of localization excluding sequence localizations; "Expression," which depicts transcriptional events and protein expression; "Companalysis," an adjunct to the sequence module for in-silico comparisons; "CV," which applies the controlled vocabularies and ontologies; "Organism," which handles species and taxonomy data; "Pub," which contains bibliographic, publications, and reference data). As depicted in Fig. 5, data currently in ChlamyDB (loci, alleles, strains, phenotypes, species, bibliographic data, genetic data, and physical maps) will be incorporated into the genetics, organism, publication, and map modules. The sequence module will be populated by nuclear, chloroplast, and mitochondrial genomic sequences, EST sequences and their assembled contigs, complete cDNA sequences obtained from our expression libraries or from information in the literature, and DNA sequences that have been used to build microarrays. In addition, the sequence module maintains relationships that link sequence records to annotation data derived from automated resources (GenBank, SwissProt, InterPro, GO, and SO, etc.) and more accurate manually curated annotation. In the future, the expression module will accommodate global gene expression data derived from the analysis of microarrays. Researchers requesting microarrays from our facility will be asked to deposit a summary of their results in this module, in addition to making their data sets publicly accessible.

(ii) Interconnecting ChlamyDB and the JGI databases. To provide a genome that has robust annotation and to avoid unnecessary duplications, ChlamyDB and the JGI will establish interdatabase links, enabling users who enter one database to retrieve data maintained by the other (Fig. 5). For example, a query of the new ChlamyDB for a particular gene or gene product will return as complete a response to the query as available and information from the JGI data set.

DOWN THE ROAD

Several important trends are emerging in *C. reinhardtii* research. Analysis of mutant phenotypes (forward genetics) will undoubtedly remain a central route for defining gene function. The availability of genomic sequence information will spur the development of insertional mutagenesis, and sequences of DNA flanking insertion sites will immediately identify putative genes responsible for specific phenotypes. Defined BAC clones will be used for rescuing mutant phenotypes, which will help establish gene function. In addition, researchers will begin to use genetic mapping of mutations on the nuclear genome to

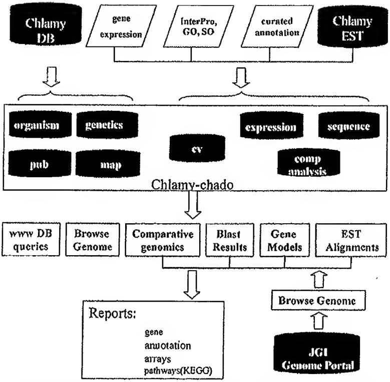


FIG. 5. Integration of databases. Data for an integrated *C. reinhardtii* database are gathered from ChlamyDB, ChlamyEST, the Chlamy database at JGI, and a variety of outside sources before being integrated in the relational database chado and served to users on the Internet. Links connecting ChlamyDB and JGI will be established to provide robust data retrieval.

routinely clone genes; one primary goal of the *C. reinhardtii* genome initiative is to provide sets of mapping primers in a 96-well format to stimulate the use of this approach. However, as genome sequence and annotation become more precise, we expect that reverse genetics will emerge as the centerpiece of functional genomics in *C. reinhardtii*, as it is now for *Arabidopsis*. This approach will exploit RNA interference and antisense RNA technologies to suppress gene expression and use tilling (74, 75, 123) to identify allelic series for specific genes; the phenotypes associated with the different alleles will help elucidate the relationship between gene structure and function.

In the very near future, global expression analyses are likely to take a central position in *C. reinhardtii* genomics. As our knowledge of transcribed regions in the genome becomes secure, construction of a full-genome microarray will be possible, enabling the synthesis of a more complete picture of the control of gene expression. Integration of the expression data will generate a catalog that describes the activity of each gene and facilitates construction of "coregulation graphs," providing clues to the physiological role of many genes of unknown function. Finally, microarray analyses applied to strains mutated for putative regulators will identify suites of genes subject to common control mechanisms.

While analysis of transcript behavior in dynamic environments will be one of the most rapid outcomes of whole genome information, many key cellular processes must be studied at the level of protein abundance and activity. The European Community is committed to building a program around *C. reinhardtii* proteomics. Initially, the focus will be to identify components localized to specific subcellular compartments, and in particular those that traffic to the chloroplast and mitochondrion. While no program currently available can accurately predict organellar targeting for *C. reinhardtii*, the results obtained by proteomic analyses should generate training sets that stimulate the development of robust predictor algorithms. Quantitative proteomics will also shape our understanding of environmental pressures that modulate levels and activities of specific proteins. Global analyses at both the protein and transcript levels, combined with computational and informatic approaches, will help predict functions of specific gene products in both metabolic and regulatory pathways and identify promoter sequences important for controlling suites of genes. Sequence information concerning promoter structure and function can be coupled with biochemical data (84, 90, 128) to determine, in a direct way, *cis*-acting sequences that modulate promoter activity. Antibodies to specific regulatory proteins identified in mutant screens can be used for chromatin immunoprecipitation (80, 88, 130), which would help establish spe-

cific protein-DNA interactions. Furthermore, two-hybrid (51, 53, 67, 124) and tandem-affinity purification (91) methodologies can be used to explore functional protein-protein interactions.

As with any organism, a strictly statistical analysis of genome sequence properties can be used to identify general and local properties of the genome such as isochores, large and small duplications, consensus sequences for splice junctions, and codon bias and its relationship to the level of expression of a gene or its evolutionary history, etc. However, because of the large underlying body of genetic, gene expression, and biochemical data, we can also predict breakthroughs in our ability to describe metabolic and regulatory pathways, and identify novel pathways as well as those that are absent or modified in specific organisms.

How *C. reinhardtii* genomics is going to evolve in the next few years is a question for the whole community. Already, the developments described here have attracted new investigators to the organism and invigorated established investigators, offering them a new pallet of tools that will undoubtedly create new landscapes in biological knowledge.

ACKNOWLEDGMENTS

We acknowledge both the National Science Foundation (for grants MCB-9975765, MCB-8819133, and DBI-9970022) and the Department of Energy for their forward-thinking support of this work.

We also thank the personnel at the Stanford Genome Center, especially Raquel Tamse, for their work in sequence the cDNAs.

A.G.W. was the Principal Investigator. All other authors in the byline are listed alphabetically.

REFERENCES

1. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389–3402.
2. Asamizu, E., K. Miura, K. Kucho, Y. Inoue, H. Fukuzawa, K. Ohyama, Y. Nakamura, and S. Tabata. 2000. Generation of expressed sequence tags from low-CO_2 and high-CO_2 adapted cells of *Chlamydomonas reinhardtii*. DNA Res. 7:305–307.
3. Asleson, C. M., and P. A. Lefebvre. 1998. Genetic analysis of flagellar length control in *Chlamydomonas reinhardtii*: a new long-flagella locus and extragenic suppressor mutations. Genetics 148:693–702.
4. Auchincloss, A. H., A. I. Loroch, and J. D. Rochaix. 1999. The argininosuccinate lyase gene of *Chlamydomonas reinhardtii*: cloning of the cDNA and its characterization as a selectable shuttle marker. Mol. Gen. Genet. 261:21–30.
5. Bateman, J. M., and S. Purton. 2000. Tools for chloroplast transformation in Chlamydomonas: expression vectors and a new dominant selectable marker. Mol. Gen. Genet. 263:404–410.
6. Bennoun, P. 1994. Chlororespiration revisited: mitochondrial-plastid interactions in *Chlamydomonas*. Biochim. Biophys. Acta 1186:59–66.
7. Bennoun, P., M. Spierer-Herz, J. Erickson, J. Girard-Bascou, Y. Pierre, M. Delosme, and J. D. Rochaix. 1986. Characterization of photosystem II mutants of *Chlamydomonas reinhardtii* lacking *psbA* gene. Plant Mol. Biol. 6:151–160.
8. Berthold, D. A., C. L. Schmidt, and R. Malkin. 1995. The deletion of *petG* in *Chlamydomonas reinhardtii* disrupts the cytochrome bf complex. J. Biol. Chem. 270:29293–29298.
9. Birney, E., and R. Durbin. 2000. Using GeneWise in the Drosophila annotation experiment. Genome Res. 10:547–548.
10. Boudreau, E., Y. Takahashi, C. Lemieux, M. Turmel, and J.-D. Rochaix. 1997. The chloroplast *ycf3* and *ycf4* open reading frames of *Chlamydomonas reinhardtii* are required for the accumulation of the photosystem I complex. EMBO J. 16:6095–6104.
11. Boudreau, E., M. Turmel, M. Goldschmidt Clermont, J. D. Rochaix, S. Sivan, A. Michaels, and S. Leu. 1997. A large open reading frame (orf 1995) in the chloroplast DNA of *Chlamydomonas reinhardtii* encodes an essential protein. Mol. Gen. Genet. 253:649–653.
12. Boynton, J. E., and N. W. Gillham. 1993. Chloroplast transformation in *Chlamydomonas*. Methods Enzymol. 217:510–536.
13. Boynton, J. E., N. W. Gillham, E. H. Harris, J. P. Hosler, A. M. Johnson, A. R. Jones, B. L. Randolph-Anderson, D. Robertson, T. M. Klein, K. B. Shark, and J. C. Sanford. 1988. Chloroplast transformation in *Chlamydomonas* with high velocity microprojectiles. Science 240:1534–1538.
14. Choquet, Y., M. Rahire, J. Girard-Bascou, J. Erickson, and J. D. Rochaix. 1992. A chloroplast gene is required for the light-independent accumulation of chlorophyll in Chlamydomonas reinhardtii. EMBO J. 11:1697–1704.
15. Cosner, M. E., R. K. Jansen, J. D. Palmer, and S. R. Downie. 1997. The highly rearranged chloroplast genome of *Trachelium caeruleum* (*Campanulaceae*): multiple inversions, inverted repeat expansion and contraction, transposition, insertions/deletions, and several repeat families. Curr. Genet. 31:419–429.
16. Davies, J., and A. R. Grossman. 1994. Sequences controlling transcription of the *Chlamydomonas reinhardtii* β_2-tubulin gene after deflagellation and during the cell cycle. Cell. Mol. Biol. 14:5165–5174.
17. Davies, J., F. Yildiz, and A. R. Grossman. 1996. *Sac1*, a putative regulator that is critical for survival of *Chlamydomonas reinhardtii* during sulfur deprivation. EMBO J. 15:2150–2159.
18. Davies, J. D., and A. R. Grossman. 1998. Responses to deficiencies in macronutrients, p. 613–635. *In* J.-D. Rochaix, M. Goldschmidt-Clermont, and S. Merchant (ed.), The molecular biology of Chlamydomonas. Kluwer Academic Publishers, Dordrecht, The Netherlands.
19. Davies, J. P., D. P. Weeks, and A. R. Grossman. 1992. Expression of the arylsulfatase gene from the β_2-tubulin promoter in *Chlamydomonas reinhardtii*. Nucleic Acids Res. 20:2959–2965.
20. Davies, J. P., F. Yildiz, and A. R. Grossman. 1994. Mutants of *Chlamydomonas reinhardtii* with aberrant responses to sulfur deprivation. Plant Cell 6:53–63.
21. Davies, J. P., F. H. Yildiz, and A. R. Grossman. 1999. *Sac3*, an Snf1-like serine/threonine kinase that positively and negatively regulates the responses of Chlamydomonas to sulfur limitation. Plant Cell 11:1179–1190.
22. Debuchy, R., S. Purton, and J. D. Rochaix. 1989. The argininosuccinate lyase gene of *Chlamydomonas reinhardtii*: an important tool for nuclear transformation and for correlating the genetic and molecular maps of the *ARG7* locus. EMBO J. 8:2803–2809.
23. Diener, D. R., A. M. Curry, K. A. Johnson, B. D. Williams, P. A. Lefebvre, K. L. Kindle, and J. L. Rosenbaum. 1990. Rescue of a paralyzed flagella mutant of *Chlamydomonas* by transformation. Proc. Natl. Acad. Sci. USA 87:5739–5743.
24. Drapier, D., J. Girard-Bascou, D. B. Stern, and F. A. Wollman. 2002. A dominant nuclear mutation in Chlamydomonas identifies a factor controlling chloroplast mRNA stability by acting on the coding region of the *atpA* transcript. Plant J. 31:687–697.
25. Drapier, D., H. Suzuki, H. Levy, B. Rimbault, K. L. Kindle, D. B. Stern, and F. A. Wollman. 1998. The chloroplast atpA gene cluster in Chlamydomonas reinhardtii. Functional analysis of a polycistronic transcription unit. Plant Physiol. 117:629–641.
26. Dutcher, S. K. 2000. *Chlamydomonas reinhardtii*: biological rationale for genomics. J. Eukaryot. Microbiol. 47:340–349.
27. Dutcher, S. K., N. S. Morrissette, A. M. Preble, C. Rackley, and J. Stanga. 2002. ε-Tubulin is an essential component of the centriole. Mol. Biol. Cell 13:3859–3869.
28. Eberhard, S., D. Drapier, and F. A. Wollman. 2002. Searching limiting steps in the expression of chloroplast-encoded proteins: relations between gene copy number, transcription, transcript abundance and translation rate in the chloroplast of *Chlamydomonas reinhardtii*. Plant J. 31:149–160.
29. Erickson, J., M. Rahire, P. Malnoe, J. Girard-Bascou, Y. Pierre, P. Bennoun, and J. D. Rochaix. 1986. Lack of D2 protein in a *Chlamydomonas reinhardtii psbD* mutant affects photosystem II stability and D1 expression. EMBO J. 5:1745–1754.
30. Fernandez, E., R. Schnell, L. P. W. Ranum, S. C. Hussey, C. D. Silflow, and P. A. Lefebvre. 1989. Isolation and characterization of the nitrate reductase structural gene of *Chlamydomonas reinhardtii*. Proc. Natl. Acad. Sci. USA 86:6449–6453.
31. Ferris, P. J., J. P. Woessner, and U. W. Goodenough. 1996. A sex recognition glycoprotein is encoded by the *plus* mating-type gene *fus1* of *Chlamydomonas reinhardtii*. Mol. Biol. Cell 7:1235–1248.
32. Finazzi, G., A. Furia, R. P. Barbagallo, and G. Forti. 1999. State transitions, cyclic and linear electron transport and photophosphorylation in *Chlamydomonas reinhardtii*. Biochim. Biophys. Acta 1413:117–129.
33. Fischer, N., E. Boudreau, M. Hippler, F. Drepper, W. Haehnel, and J. D. Rochaix. 1999. A large fraction of PsaF is nonfunctional in photosystem I complexes lacking the PsaJ subunit. Biochemistry 38:5546–5552.
34. Fischer, N., P. Setif, and J.-D. Rochaix. 1997. Targeted mutations in the *psaC* gene of *Chlamydomonas reinhardtii*: preferential reduction of FB at low temperature is not accompanied by altered electron flow from photosystem I to ferredoxin. Biochemistry 36:93–102.
35. Fischer, N., O. Stampacchia, K. Redding, and J. D. Rochaix. 1996. Selectable marker recycling in the chloroplast. Mol. Gen. Genet. 251:373–380.
36. Franklin, S., B. Ngo, E. Efuet, and S. P. Mayfield. 2002. Development of a GFP reporter gene for *Chlamydomonas reinhardtii* chloroplast. Plant J. 30:733–744.
37. Fuhrmann, M., W. Oertel, and P. Hegemann. 1999. A synthetic gene coding

for the green fluorescent protein (GFP) is a versatile reporter in *Chlamydomonas reinhardtii*. Plant J. 19:353–361.
38. Funke, R. P., J. L. Kovar, and D. P. Weeks. 1997. Intracellular carbonic anhydrase is essential to photosynthesis in *Chlamydomonas reinhardtii* at atmospheric levels of CO_2. Demonstration via genomic complementation of the high-CO_2-requiring mutant *ca-1*. Plant Physiol. 114:237–244.
39. Goldschmidt-Clermont, M. 1991. Transgenic expression of aminoglycoside adenine transferase in the chloroplast: a selectable marker for site-directed transformation of Chlamydomonas. Nucleic Acids Res. 19:4083–4089.
40. Goldschmidt-Clermont, M., Y. Choquet, J. Girard-Bascou, F. Michel, M. Schirmer-Rahire, and J. Rochaix. D. 1991. A small chloroplast RNA may be required for trans-splicing in *Chlamydomonas reinhardtii*. Cell 65:135–144.
41. Gross, C. H., L. P. Ranum, and P. A. Lefebvre. 1988. Extensive restriction fragment length polymorphisms in a new isolate of *Chlamydomonas reinhardtii*. Curr. Genet. 13:503–508.
42. Grossman, A. 2000. *Chlamydomonas reinhardtii* and photosynthesis: genetics to genomics. Curr. Opin. Plant Biol. 3:132–137.
43. Hallahan, B. J., S. Purton, A. Ivison, D. Wright, and M. C. W. Evans. 1995. Analysis of the proposed Fe-Sx binding region in *Chlamydomonas reinhardtii*. Photosynth. Res. 46:257–264.
44. Harris, E. H. 2001. Chlamydomonas as a model organism. Annu. Rev. Plant Physiol. Plant Mol. Biol. 52:363–406.
45. Higgs, D. C., R. S. Shapiro, K. L. Kindle, and D. B. Stern. 1999. Small cis acting sequences that specify secondary structures in a chloroplast mRNA are essential for RNA stability and translation. Mol. Cell Biol. 19:8479–8491.
46. Hong, S., and R. J. Spreitzer. 1997. Complementing substitutions at the bottom of the barrel influence catalysis and stability of ribulose-bisphosphate carboxylase-oxygenase. J. Biol. Chem. 272:11114–11117.
47. Huang, C., S. Wang, L. Chen, C. Lemieux, C. Otis, M. Turmel, and X. Q. Liu. 1994. The Chlamydomonas chloroplast *clpP* gene contains translated large insertion sequences and is essential for cell growth. Mol. Gen. Genet. 244:151–159.
48. Im, C. S., Z. Zhang, J. Shrager, C. W. Chang, and A. R. Grossman. 2003. Analysis of light and CO_2 regulation in *Chlamydomonas reinhardtii* using genome-wide approaches. Photosynth. Res. 75:111–125.
49. Jacobshagen, S., K. L. Kindle, and C. H. Johnson. 1996. Transcription of CABII is regulated by the biological clock in *Chlamydomonas reinhardtii*. Plant Mol. Biol. 31:1173–1184.
50. Jeong, B. R., D. Wu-Scharf, C. Zhang, and H. Cerutti. 2002. Suppressors of transcriptional transgenic silencing in Chlamydomonas are sensitive to DNA-damaging agents and reactivate transposable elements. Proc. Natl. Acad. Sci. USA 99:1076–1081.
51. Joung, J. K. 2001. Identifying and modifying protein-DNA and protein-protein interactions using a bacterial two-hybrid selection system. J Cell Biochem Suppl. Suppl. 37:53–57.
52. Kang, T. J., and M. W. Fawley. 1997. Variable (CA/GT)n simple sequence repeat DNA in the alga Chlamydomonas. Plant Mol. Biol. 35:943–948.
53. Karimova, G., D. Ladant, and A. Ullmann. 2002. Two-hybrid systems and their usage in infection biology. Int. J. Med. Microbiol. 292:17–25.
54. Karolchik, D., R. Baertsch, M. Diekhans, T. S. Furey, A. Hinrichs, Y. T. Lu, K. M. Roskin, M. Schwartz, C. W. Sugnet, D. J. Thomas, R. J. Weber, D. Haussler, and W. J. Kent. 2003. The UCSC Genome Browser Database. Nucleic Acids Res. 31:51–54.
55. Kathir, P., M. LaVoie, W. J. Brazelton, N. A. Haas, P. A. Lefebvre, and C. D. Silflow. 2003. Molecular map of the *Chlamydomonas reinhardtii* nuclear genome. Eukaryot. Cell. 2:362–379.
56. Kindle, K. L. 1990. High-frequency nuclear transformation of *Chlamydomonas reinhardtii*. Proc. Natl. Acad. Sci. USA 87:1228–1232.
57. Kindle, K. L., R. A. Schnell, E. Fernández, and P. A. Lefebvre. 1989. Stable nuclear transformation of *Chlamydomonas* using the *Chlamydomonas* gene for nitrate reductase. J. Cell Biol. 109:2589–2601.
58. Koblenz, B., J. Schoppmeier, A. Gronow, and K. F. Lechtreck. 2003. Centrin deficiency in Chlamydomonas causes defects in basal body replication, segregation and maturation. J. Cell Sci. 116:2635–2646.
59. Komine, Y., E. Kikis, G. Schuster, and D. Stern. 2002. Evidence for in vivo modulation of chloroplast RNA stability by 3'-UTR homopolymeric tails in *Chlamydomonas reinhardtii*. Proc. Natl. Acad. Sci. USA 99:4085–4090.
60. Koutoulis, A., G. J. Pazour, C. G. Wilkerson, K. Inaba, H. Sheng, S. Takada, and G. B. Witman. 1997. The *Chlamydomonas reinhardtii ODA3* gene encodes a protein of the outer dynein arm docking complex. J. Cell Biol. 137:1069–1080.
61. Kuchka, M. R., ... — 61. Kunstner, P., A. Guardiola, Y. Takahashi, and J. D. Rochaix. 1995. A mutant strain of *Chlamydomonas reinhardtii* lacking the chloroplast photosystem II psbI gene grows photoautotrophically. J. Biol. Chem. 270:9651–9654.
62. Kuras, R., and F.-A. Wollman. 1994. The assembly of cytochrome b6/f complexes: an approach using genetic transformation of the green alga *Chlamydomonas reinhardtii*. EMBO J. 13:1019–1027.
63. Lardans, A., N. W. Gillham, and J. E. Boynton. 1997. Site-directed mutations at residue 251 of the photosystem II D1 protein of *Chlamydomonas* that result in a nonphotosynthetic phenotype and impair D1 synthesis and accumulation. J. Biol. Chem. 272:210–216.
64. Larson, E. M., C. M. O'Brien, G. Zhu, R. J. Spreitzer, and A. R. Portis, Jr. 1997. Specificity for activase is changed by a Pro-89 to Arg substitution in the large subunit of ribulose-1, 5-bisphosphate carboxylase-oxygenase. J. Biol. Chem. 272:17033–17037.
65. Lechtreck, K. F., J. Rostmann, and A. Grunow. 2002. Analysis of Chlamydomonas SF-assemblin by GFP tagging and expression of antisense constructs. J. Cell Sci. 115:1511–1522.
66. Lefebvre, P. A., and C. D. Silflow. 1999. Chlamydomonas: the cell and its genomes. Genetics 151:9–14.
67. Legrain, P., J. Wojcik, and J. M. Gauthier. 2001. Protein-protein interaction maps: a lead towards cellular functions. Trends Genet. 17:346–352.
68. Lilly, J. W., J. E. Maul, and D. B. Stern. 2002. The *Chlamydomonas reinhardtii* organellar genomes respond transcriptionally and post-transcriptionally to abiotic stimuli. Plant Cell 14:2681–2706.
69. Liu, X.-Q., C. Huang, and H. Xu. 1993. The unusual *rps3*-like orf712 is functionally essential and structurally conserved in *Chlamydomonas*. FEBS Lett. 336:225–230.
70. Lumbreras, V., D. R. Stevens, and S. Purton. 1998. Efficient foreign gene expression in *Chlamydomonas reinhardtii* mediated by an endogenous intron. Plant J. 14:441–447.
71. Majeran, W., F. A. Wollman, and O. Vallon. 2000. Evidence for a role of ClpP in the degradation of the chloroplast cytochrome b(6)f complex. Plant Cell. 12:137–150.
72. Maul, J. E., J. W. Lilly, L. Cui, C. W. dePamphilis, W. Miller, E. H. Harris, and D. B. Stern. 2002. The *Chlamydomonas reinhardtii* plastid chromosome: islands of genes in a sea of repeats. Plant Cell. 14:2659–2679.
73. Mayfield, S. P., and K. L. Kindle. 1990. Stable nuclear transformation of *Chlamydomonas reinhardtii* by using a *C. reinhardtii* gene as the selectable marker. Proc. Natl. Acad. Sci. USA 87:2087–2091.
74. McCallum, C. M., L. Comai, E. A. Greene, and S. Henikoff. 2000. Targeted screening for induced mutations. Nat. Biotechnol. 18:455–457.
75. McCallum, C. M., L. Comai, E. A. Greene, and S. Henikoff. 2000. Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. Plant Physiol. 123:439–442.
76. Melkozernov, A. N., H. Su, S. Lin, S. Bingham, A. N. Webber, and R. E. Blankenship. 1997. Specific mutations near the primary donor in photosystem I from *Chlamydomonas reinhardtii* alters the trapping time and spectroscopic properties of P700. Biochemistry 36:2898–2907.
77. Minko, I., S. P. Holloway, S. Nikaido, M. Carter, O. W. Odom, C. H. Johnson, and D. L. Herrin. 1999. Renilla luciferase as a vital reporter for chloroplast gene expression in Chlamydomonas. Mol. Gen. Genet. 262:421–425.
78. Morais, F., J. Barber, and P. J. Nixon. 1998. The chloroplast-encoded alpha subunit of cytochrome b-559 is required for assembly of the photosystem two complex in both the light and the dark in *Chlamydomonas reinhardtii*. J. Biol. Chem. 273:29315–29320.
79. Mulder, N. J., R. Apweiler, T. K. Attwood, A. Bairoch, D. Barrell, A. Bateman, D. Binns, M. Biswas, P. Bradley, P. Bork, P. Bucher, R. R. Copley, E. Courcelle, U. Das, R. Durbin, L. Falquet, W. Fleischmann, S. Griffiths-Jones, D. Haft, N. Harte, N. Hulo, D. Kahn, A. Kanapin, M. Krestyaninova, R. Lopez, I. Letunic, D. Lonsdale, V. Silventoinen, S. E. Orchard, M. Pagni, D. Peyruc, C. P. Ponting, J. D. Selengut, F. Servant, C. J. Sigrist, R. Vaughan, and E. M. Zdobnov. 2003. The InterPro Database, 2003 brings increased coverage and new features. Nucleic Acids Res. 31:315–318.
80. Nal, B., E. Mohr, and P. Ferrier. 2001. Location analysis of DNA-bound proteins at the whole-genome level: untangling transcriptional regulatory networks. Bioessays 23:473–476.
81. Nelson, J. A. E., and P. A. Lefebvre. 1995. Targeted disruption of the *NIT8* gene in *Chlamydomonas reinhardtii*. Mol. Cell. Biol. 15:5762–5769.
82. Nelson, J. A. E., P. B. Savereide, and P. A. Lefebvre. 1994. The *CRY1* gene in *Chlamydomonas reinhardtii*: structure and use as a dominant selectable marker for nuclear transformation. Mol. Cell. Biol. 14:4011–4019.
83. Newman, S. M., J. E. Boynton, N. W. Gillham, B. L. Randolph-Anderson, A. M. Johnson, and E. H. Harris. 1990. Transformation of chloroplast ribosomal RNA in *Chlamydomonas*: molecular and genetic characterization of integration events. Genetics 126:875–888.
84. Nielsen, P. E. 1989. In vivo footprinting: studies of protein-DNA interactions in gene regulation. Bioessays 11:152–155.
85. O'Connor, H. E., S. V. Ruffle, A. J. Cain, Z. Deak, I. Vass, J. H. Nugent, and S. Purton. 1998. The 9-kDa phosphoprotein of photosystem II. Generation and characterisation of Chlamydomonas mutants lacking PSII-H and a site-directed mutant lacking the phosphorylation site. Biochim. Biophys. Acta 1364:63–72.
86. Ohnishi, N., and Y. Takahashi. 2001. PsbT polypeptide is required for efficient repair of photodamaged photosystem II reaction center. J. Biol. Chem. 276:33798–33804.
87. Ohresser, M., R. F. Matagne, and R. Loppes. 1997. Expression of the arylsulphatase reporter gene under the control of the *NIT1* promoter of *Chlamydomonas reinhardtii*. Curr. Genet. 31:264–271.

88. Orlando, V. 2000. Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation. Trends Biochem. Sci. 25:99–104.
89. Palombella, A. L., and S. K. Dutcher. 1998. Identification of the gene encoding the tryptophan synthase beta-subunit from *Chlamydomonas reinhardtii*. Plant Physiol. 117:455–464.
90. Pfeifer, G. P., and S. Tornaletti. 2000. In vivo footprinting using UV light and ligation-mediated PCR. Methods Mol. Biol. 130:13–27.
91. Puig, O., F. Caspary, G. Rigaut, B. Rutz, E. Bouveret, E. Bragado-Nilsson, M. Wilm, and B. Seraphin. 2001. The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods 24:218–229.
92. Purton, S., and J.-D. Rochaix. 1994. Complementation of a *Chlamydomonas reinhardtii* mutant using a genomic cosmid library. Plant Mol. Biol. 24:533–537.
93. Quinn, J. M., and S. Merchant. 1995. Two Copper-responsive elements associated with the *Chlamydomonas Cyc6* gene function as targets for transcriptional activators. Plant Cell 7:623–638.
94. Randolph-Anderson, B. L., R. Sato, A. M. Johnson, E. H. Harris, C. R. Hauser, K. Oeda, F. Ishige, S. Nishio, N. W. Gillham, and J. E. Boynton. 1998. Isolation and characterization of a mutant protoporphyrinogen oxidase gene from *Chlamydomonas reinhardtii* conferring resistance to porphyric herbicides. Plant Mol. Biol. 38:839–859.
95. Redding, K., L. Cournac, I. R. Vassiliev, J. H. Golbeck, G. Peltier, and J. D. Rochaix. 1999. Photosystem I is indispensable for photoautotrophic growth, CO_2 fixation, and H_2 photoproduction in *Chlamydomonas reinhardtii*. J. Biol. Chem. 274:10466–10473.
96. Robertson, D., J. E. Boynton, and N. W. Gillham. 1990. Cotranscription of the wild-type chloroplast atpE gene encoding the CF1/CF0 epsilon subunit with the 3′ half of the rps7 gene in Chlamydomonas reinhardtii and characterization of frameshift mutations in atpE. Mol. Gen. Genet. 221:155–163.
97. Rochaix, J. D. 2002. Chlamydomonas, a model system for studying the assembly and dynamics of photosynthetic complexes. FEBS Lett. 529:34–38.
98. Rochaix, J.-D. 1995. *Chlamydomonas reinhardtii* as the photosynthetic yeast. Annu. Rev. Genet. 29:209–230.
99. Rochaix, J.-D. 1996. Post-transcriptional regulation of chloroplast gene expression in *Chlamydomonas*. Plant Mol. Biol. 32:327–341.
100. Rochaix, J.-D., M. Kuchka, S. Mayfield, M. Schirmer-Rahire, J. Girard-Bascou, and P. Bennoun. 1989. Nuclear and chloroplast mutations affect the synthesis or stability of the chloroplast *psbC* gene product in *Chlamydomonas reinhardtii*. EMBO J. 8:1013–1021.
101. Rolland, N., A. J. Dorne, G. Amoroso, D. F. Sultemeyer, J. Joyard, and J. D. Rochaix. 1997. Disruption of the plastid *ycf10* open reading frame affects uptake of inorganic carbon in the chloroplast of *Chlamydomonas*. EMBO J. 16:6713–6726.
102. Schnell, R. A., and P. A. Lefebvre. 1993. Isolation of the Chlamydomonas regulatory gene *NIT2* by transposon tagging. Genetics 134:737–747.
103. Schroda, M., O. Vallon, F. A. Wollman, and C. F. Beck. 1999. A chloroplast-targeted heat shock protein 70 (HSP70) contributes to the photoprotection and repair of photosystem II during and after photoinhibition. Plant Cell. 11:1165–1178.
104. Shepherd, H. S., J. E. Boynton, and N. W. Gillham. 1979. Mutations in nine chloroplast loci of *Chlamydomonas* affecting different photosynthetic functions. Proc. Natl. Acad. Sci. USA 76:1353–1357.
105. Shimogawara, K., S. Fujiwara, A. R. Grossman, and H. Usuda. 1998. High efficiency transformation of *Chlamydomonas reinhardtii* by electroporation. Genetics 148:1821–1828.
106. Shrager, J., C. Hauser, C. W. Chang, E. H. Harris, J. Davies, J. McDermott, R. Tamse, Z. Zhang, and A. R. Grossman. 2003. *Chlamydomonas reinhardtii* genome project. A guide to the generation and use of the cDNA information. Plant Physiol. 131:401–408.
107. Sineshchekov, O. A., K. H. Jung, and J. L. Spudich. 2002. Two rhodopsins mediate phototaxis to low- and high-intensity light in *Chlamydomonas reinhardtii*. Proc. Natl. Acad. Sci. USA 99:225–230.
108. Smith, E. F., and P. A. Lefebvre. 1996. PF16 encodes a protein with armadillo repeats and localizes to a single microtubule of the central apparatus in Chlamydomonas flagella. J. Cell Biol. 132:359–370.
109. Smith, E. F., and P. A. Lefebvre. 1997. PF20 gene product contains WD repeats and localizes to the intermicrotubule bridges in Chlamydomonas flagella. Mol. Biol. Cell 8:455–467.
110. Spreitzer, R. J., M. Goldschmidt-Clermont, M. Rahire, and J. D. Rochaix. 1985. Nonsense mutations in the Chlamydomonas chloroplast gene that codes for the large subunit of ribulose bisphosphate carboxylase/oxygenase. Proc. Natl. Acad. Sci. USA 82:5460–5464.
111. Stein, L. D., C. Mungall, S. Shu, M. Caudy, M. Mangone, A. Day, E. Nickerson, J. E. Stajich, T. W. Harris, A. Arva, and S. Lewis. 2002. The generic genome browser: a building block for a model organism system database. Genome Res. 12:1599–1610.
112. Stevens, D. R., J.-D. Rochaix, and S. Purton. 1996. The bacterial phleomycin resistance gene *ble* as a dominant selectable marker in *Chlamydomonas*. Mol. Gen. Genet. 251:23–30.
113. Summer, E. J., V. H. Schmid, B. U. Bruns, and G. W. Schmidt. 1997. Requirement for the H phosphoprotein in photosystem II of Chlamydomonas reinhardtii. Plant Physiol. 113:1359–1368.
114. Suzuki, J. Y., and C. E. Bauer. 1992. Light-independent chlorophyll biosynthesis: involvement of the chloroplast gene chlL (frxC). Plant Cell 4:929–940.
115. Swiatek, M., R. Kuras, A. Sokolenko, D. Higgs, J. Olive, G. Cinque, B. Muller, L. A. Eichacker, D. B. Stern, R. Bassi, and R. G. Herrmann. 2001. The chloroplast gene *ycf9* encodes a photosystem II (PSII) core subunit, PsbZ, that participates in PSII supramolecular structure. Plant Cell 13:1347–1367.
116. Takahashi, H., C. E. Braby, and A. R. Grossman. 2001. Sulfur economy and cell wall biosynthesis during sulfur limitation of *Chlamydomonas reinhardtii* Plant Physiol. 127:665–673.
117. Takahashi, Y., M. Goldschmidt-Clermont, S. S.-Y., L. G. Franzen, and J.-D. Rochaix. 1991. Directed chloroplast transformation in *Chlamydomonas reinhardtii*: insertional inactivation of the *psaC* gene encoding the iron-sulfur protein destabilizes photosystem I. EMBO J. 10:2033–2040.
118. Takahashi, Y., H. Matsumoto, M. Goldschmidt-Clermont, and J.-D. Rochaix. 1994. Directed disruption of the *Chlamydomonas* chloroplast *psbK* gene destabilizes the photosystem II reaction center complex. Plant Mol Biol. 24:779–788.
119. Takahashi, Y., M. Rahire, C. Breyton, J. L. Popot, P. Joliot, and J. D. Rochaix. 1996. The chloroplast ycf7 (petL) open reading frame of Chlamydomonas reinhardtii encodes a small functionally important subunit of the cytochrome b6f complex. EMBO J. 15:3498–3506.
120. Tam, L.-W., and P. A. Lefebvre. 1993. Cloning of flagellar genes in *Chlamydomonas reinhardtii* by DNA insertional mutagenesis. Genetics 135:375–384.
121. Tang, D. K. H., S.-Y. Qiao, and M. Wu. 1995. Insertion mutagenesis of *Chlamydomonas reinhardtii* by electroporation and heterologous DNA. Biochem. Mol. Biol. Int. 36:1025–1035.
122. Thompson, R. J., and G. Mosig. 1990. Light affects the structure of Chlamydomonas chloroplast chromosomes. Nucleic Acids Res. 18:2625–2631.
123. Till, B. J., S. H. Reynolds, E. A. Greene, C. A. Codomo, L. C. Enns, J. E. Johnson, C. Burtner, A. R. Odden, K. Young, N. E. Taylor, J. G. Henikoff, L. Comai, and S. Henikoff. 2003. Large-scale discovery of induced point mutations with high-throughput TILLING. Genome Res. 13:524–530.
124. Toby, G. G., and E. A. Golemis. 2001. Using the yeast interaction trap and other two-hybrid-based approaches to study protein-protein interactions. Methods 24:201–217.
125. Villand, P., M. Eriksson, and G. Samuelsson. 1997. Regulation of genes by the environmental CO_2 level. Plant Physiol. 114:258–259.
126. Vysotskaia, V. S., D. E. Curtis, A. V. Voinov, P. Kathir, C. D. Silflow, and P. A. Lefebvre. 2001. Development and characterization of genome-wide single nucleotide polymorphism markers in the green alga *Chlamydomonas reinhardtii*. Plant Physiol. 127:386–389.
127. Wang, S.-C., R. A. Schnell, and P. A. Lefebvre. 1998. Isolation and characterization of a new transposable element in *Chlamydomonas reinhardtii*. Plant Mol. Biol. 38:681–687.
128. Warshawsky, D., and L. Miller. 1999. Mapping protein-DNA interactions using in vivo footprinting. Methods Mol. Biol. 127:199–212.
129. Webber, A. N., H. Su, S. E. Bingham, H. Kass, L. Krabben, M. Kuhn, E. Schlodder, and W. Lubitz. 1996. Site-directed mutations affecting the spectroscopic characteristics and mid-point potential of the primary donor in photosystem I. Biochemistry 39:12857–12863.
130. Weinmann, A. S., and P. J. Farnham. 2002. Identification of unknown target genes of human transcription factors using chromatin immunoprecipitation. Methods 26:37–47.
131. Whitelegge, J. P., D. Koo, and J. Erickson. 1992. Site-directed mutagenesis of the chloropolast *psbA* gene encoding the D1 polypeptide of photosystem II in *Chlamydomonas reinhardtii* changes at aspartate 170 affect the assembly of a functional water-splitting manganese cluster, p. 151–154. *In* N. Murata (ed.), Research in photosynthesis, vol. II. Kluwer Academic Publishers, Dordrecht, The Netherlands.
132. Wykoff, D., A. Grossman, D. P. Weeks, H. Usuda, and K. Shimogawara. 1999. Psr1, a nuclear localized protein that regulates phosphorus metabolism in *Chlamydomonas*. Proc. Natl. Acad. Sci. USA 96:15336–15341.
133. Xie, Z., and S. Merchant. 1996. The plastid-encoded *ccsA* gene is required for heme attachment to chloroplast c-type cytochromes. J. Biol. Chem. 271:4632–4639.
134. Xiong, J., R. S. Hutchinson, R. T. Sayre, and Govindjee. 1997. Modification of the photosystem II acceptor side function in a D1 mutant (arginine-269-glycine) of *Chlamydomonas reinhardtii*. Biochim. Biophys. Acta 1322:60–76.
135. Yehudai-Resheff, S., M. Hirsh, and G. Schuster. 2001. Polynucleotide phosphorylase functions as both an exonuclease and a poly(A) polymerase in spinach chloroplasts. Mol. Cell. Biol. 21:5408–5416.
136. Yildiz, F. H., J. P. Davies, and A. R. Grossman. 1996. Sulfur availability and the *SAC1* gene control adenosine triphosphate sulfurylase gene expression in *Chlamydomonas reinhardtii*. Plant Physiol. 112:669–675.

137. Zerges, W., A. H. Auchincloss, and J. D. Rochaix. 2003. Multiple translational control sequences in the 5' leader of the chloroplast *psbC* mRNA interact with nuclear gene products in *Chlamydomonas reinhardtii*. Genetics 163:895–904.
138. Zhang, D., and P. A. Lefebvre. 1997. FAR1, a negative regulatory locus required for the repression if the nitrate reductase gene in *Chlamydomonas reinhardtii*. Genetics 146:121–133.
139. Zhang, H., P. L. Herman, and D. P. Weeks. 1994. Gene isolation through genomic complementation using an indexed library of *Chlamydomonas reinhardtii* DNA. Plant Mol Biol. 24:663–672.
140. Zhu, G., and R. J. Spreitzer. 1996. Directed mutagenesis of chloroplast ribulose-1, 5-bisphosphate carboxylase-oxygenase. Loop 6 substitutions complement for structural stability but decrease catalytic efficiency. J. Biol. Chem. 271:18494–18498.